LONG-TERM IMPACTS OF WATER QUALITY AND CLIMATE VARIABILITY ON ILLINOIS RIVER WATERWAY FISH ASSEMBLAGES

BY

JERROD LEE PARKER

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Natural Resources and Environmental Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Master's Committee:

Adjunct Associate Professor Yong Cao, Co-adviser
Adjunct Professor John Epifanio, Co-adviser
Assistant Professor Michael P. Ward
Associate Professor Rebecca C. Fuller

Abstract

The Illinois River has been heavily altered by a range of human activities. Among these, wetlands have been drained, lock and dams constructed, navigation channels maintained, and industrial effluent and sewage discharged. These alterations have strongly affected fish assemblages within the river. In 1957, the Illinois Natural History Survey began the Long-Term Fish Population Monitoring Program to sample fish assemblages throughout the Illinois River Waterway (IRW) on a yearly basis. During the 1960s, researchers observed that much of the upper river was depauperate. Fishes consisted almost entirely of pollution tolerant non-native common carp (*Cyprinus carpio*) and goldfish (*Carassius auratus*). Following implementation of the 1972 Clean Water Act (CWA), the upper river fish assemblages gradually became much more diverse. This increased diversity was thought to have resulted from pollution mitigation efforts, mandated by the CWA. However, the relationships between fish assemblages and specific water quality variables have not been established. Additionally, interannual climate variability may have significantly affected flow, water temperature, and water quality that, in turn, affected fish assemblages. My goal was to redress the lack of statistical support for the role of water quality and climate variability in multiple descriptors of fish assemblages.

I obtained comprehensive water quality data from the Metropolitan Water Reclamation District of Greater Chicago, and regional climate data from the National Climatic Data Center for the upper 207km of the IRW between 1983 and 2010. I used multiple linear and random forests regression to assess the relative importance of water quality and climate variables on multiple fish assemblage attributes (e.g., proportion of predators, species richness, and functional diversity). Models found ammonia, dissolved oxygen, and clarity most capable of explaining the observed assemblage changes. Climate played a smaller, but important role. Moreover, assemblage attributes based on fish biomass tended to be better explained by the environmental data than those based on fish abundance. My results indicate that, even in a heavily modified and regulated river, better water quality management alone can substantially improve the biodiversity and structure of fish assemblages. The results also highlight specific water quality variables (e.g. ammonia, dissolved oxygen, clarity) to direct the focus of monitoring and regulation programs.

Acknowledgements

During my career I have enjoyed the support, advisement, and mentoring of several very influential individuals. First, Dr. Wayne Webber who introduced me to academic research. Second, Dr. Kristopher Wright who helped me discover my chosen field of study, by introducing me to the field of freshwater ecology. Third, Jean Unmuth who allowed me the opportunity to work with her, and took time to answer every question I had. Fourth, Dave Marshall who inspired my appreciation for the ecological complexity of large rivers. Finally, Dr. Yong Cao and Dr. John Epifanio whom have given invaluable writing and statistical guidance, and supported my graduate endeavors.

This project was only possible due to the funding support of The Illinois Department of Natural Resources and the United States Fish and Wildlife Service Federal Aid in Sportfish Restoration Project F-101-R. Thank you to Dr. Greg Sass for allowing me the opportunity to work with such a historically significant dataset. Thanks to my committee, Doctors Becky Fuller and Mike Ward, for helping me flesh out my ideas in the developmental stages. I am grateful to everyone at the Illinois River Biological Station that has collected data throughout the years, and Dr. Richard E. Sparks who helped me develop my approach with his numerous published works and conversation. Finally, a special thank you to the Metropolitan Water Reclamation District and Jennifer Wasik for allowing me the use of a great dataset, and promptly answering a multitude of questions.

Table of Contents

Chapter 1

Introduction and Literature Review

Ecosystem Services of Large River Ecosystems

Throughout history, humans have depended upon the ecosystem services provided by large river ecosystems. Historians point to civilization beginning in river valleys across the globe. The Nile, Tigris, Yellow, Indus, and other rivers birthed great civilizations (Moigne and Priscoli, 1992) . Much of modern society remains dependent upon services provided by rivers for fresh water, fertile soil, food sources, transportation corridors, hydropower, recreation, and a general sense of well-being. Evidence of this dependence can be seen globally, as rivers are the nearest freshwater feature for the majority of the human population and population densities are highest near large rivers (Kummu et al., 2011).

Rivers provide freshwater necessary to human and domesticated plant and animal life. Agriculture accounts for approximately 70% of water withdrawals from river basins globally (Smakhtin et al., 2004). Most people in the United States (approximately 195 million) obtain their drinking water from surface sources, with rivers directly supplying many large cities (e.g. Seattle, New York, Philadelphia, Los Angeles, Las Vegas, Atlanta) (Harris, 2013; Kratzer et al., 2006; Petes et al., 2012; Richards et al., 2012; Urbansky et al., 2000). Further, many groundwater sources are dependent upon large river basins for recharge (Liang et al., 2003; Vaughn and Hakenkamp, 2001; Vought et al., 1994).

Recreation is another economically and societally important benefit of rivers (Bockstael et al., 2000). Between 1995 and 2008 mean annual United States duck stamp sales were $22,892,820 (Vrtiska et al., 2013). Much of this hunting takes place in or near large rivers, migration corridors for waterfowl (Sparks, 1995). The 2011 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation reports freshwater fishermen in the United States spent an estimated $25.7 billion on trip and equipment expenses alone, or roughly double the amount spent on saltwater fishing, $10.3 billion (USFWS, 2011). The activities of sportsman not only result in food harvests, but also in an unquantifiable sense of well-being acquired through their activities (Kozicky, 1977). Related to this sense of well-being and enjoyment are activities that take advantage of rivers directly, or the trails and parks often placed along rivers (e.g. wildlife-watching, hiking, biking, recreational boating).

Rivers also provide ready-made transportation corridors. These highways are vital for the exchange of goods, information, and disposal of wastes. Early on in human development, rivers were essential in the dispersal of peoples, goods, and ideas (Church and Bell, 1988). Modern technology has replaced rivers for the movement of peoples and ideas, but rivers are still important for the movement of goods and wastes. A recent analysis of long haul shipping costs found barge transport to be around a tenth of the cost of rail or semi (Gonzales et al., 2013). Industrial wastewater use of, and inputs into rivers has been increasing globally (Saravanan et al., 2011). The growing population also relies upon rivers to transport sewage away from communities. In the United States combined sewer systems serve approximately 40 million people, with the majority passing through rivers before entering oceans (Aronson et al., 2012).

The energy of river systems has been harnessed since antiquity. Use of water wheels to power grain mills dates back to between 200 to 100 years Before the Common Era. (Avitsur, 1960). With the threat of fossil fuel emission induced climate change, hydroelectric power is becoming increasingly important in modern times. In the United States, it accounts for 67% of domestic renewable energy, or around 7% of total electricity generated (O'Connor, 2013). Globally hydropower is more important, accounting for almost 17% of total energy production in 2010 (Zimny et al., 2013).

The biological productivity of river systems is another function that greatly benefits human society. Alluvial river valleys provide extremely fertile soils, which, if uninterrupted, are replenished annually through seasonal flooding (Banerjee, 2010; Schroeder, 1999). River-floodplain ecosystems are among the most productive ecosystems in the world (Bayley, 1995). Some of the functions floodplains facilitate are groundwater recharge, water purification, nutrient cycling, flood relief, and the maintenance of fish populations. The importance of floodplain lakes to fish assemblages was quickly recognized when Stephen A. Forbes published *The Lake as a Microcosm* in 1887 (Forbes, 1925). Floodplains are particularly important to North American game fish species, which utilize these areas as overwintering and spawning locations. They enhance river populations of such fishes as crappie, bluegill, bass, and pike (Kwak, 1988; Ross and Baker, 1983).

In this study, I focus on fish due to their importance to humanity as a source of protein, recreation, and to overall river health. The total commercial fish catch of inland waters worldwide in 2002 was 8.7 million tons, with developing countries accounting for the majority of catches (Allan et al., 2005). Though fish harvests are depended upon less in developed countries, commercial harvests in the Upper Mississippi River (UMR) in 1998 was approximately 5 million kilograms with a value just under 4 million dollars (Schramm Jr, 2004). In the Pacific Northwest United States salmon fishing in 1995 supported an estimated 60,000 jobs equivalent to over $1 billion dollars in income (Brown and Shogren, 1998).

Fish also mediate many ecosystem functions, which are difficult to quantify economically. For example, fish are important to sediment processes, carbon and nutrient flux, and linking rivers with terrestrial and marine environments (Holmlund and Hammer, 1999). Multiple benthic fish species are known to resuspend sediment and organic matter through foraging behaviors which effects water clarity, and carbon and nutrient distribution (Breukelaar et al., 1994; Gelwick et al., 1997). Fish consumption of terrestrial food sources act as an energetic link between land and rivers, and migration of anadromous fishes transports marine nutrients to river systems. Additionally, the upper trophic levels fish occupy allow for assessments of pollution (e.g. xenobiotics, mercury) through bioaccumulation (Mason et al., 2000; Van der Oost et al., 2003).

River Alterations Affecting Fish Biodiversity

Many human activities have affected large river ecosystem health, particularly fish biodiversity. Biodiversity increases the efficiency of biological functions, and allows for functional redundancies that enable function to be maintained and recover from perturbations (Cardinale et al., 2006; Chapin III et al., 1997; Srivastava and Vellend, 2005). Both efficiency and the ability to adapt to perturbations create a stable, sustainable, resilient, and autonomous ecosystem able to provide and maintain biologically derived ecosystem services. The positive relationship between biodiversity and ecosystem services in rivers was reported by Nelson et al. (2009), as biodiversity can enhance water quality, soil retention, and carbon sequestration.

Water quality is a primary stressor for fish assemblages. Several aspects of water quality are impacted by human pollution, hydrological modifications, and the interaction between the two. Affected parameters include sediment loads, nutrients, metals, temperature, pathogens, and a broad spectrum of organic compounds. Each of these aspects of water quality has individual and interacting effects on fish assemblages (Meybeck, 2003; Meybeck, 2004).

Anthropogenic effects on sediment loads are complex, and have extensive ecological consequences. Land use changes (e.g. forest removal, mining, quarrying, construction) in the late nineteenth and early twentieth century globally increased suspended sediment loads in large rivers an average of 3.5 times pre-industrialization levels (Douglas, 1990). North America alone lost an estimated 3,600 million tons of soil annually during the 1930's, of which 75% came directly from cropland (Douglas, 1990). Strikingly, this has not resulted in significant sediment transport increases at river mouths, mainly due to the efficiency with which dams and floodplains trap sediments.(Meybeck, 2003). Meybeck (2003) estimated reservoirs currently store more than 30% of global river sediment loads.

At the local level, increased suspended sediments alter water clarity by increasing turbidity. This lowers and alters wavelengths of light penetrating the water column, thus changes the way primary producer communities are structured (Bunn and Arthington, 2002). Changes in primary production can cause bottom up trophic cascades, which alter fish assemblages. Moreover, decreases in water clarity adversely affect fish dependent on sight for predation and predator avoidance (Brusven and Rose, 1981; Johnson and Hines, 1999; Shingles et al., 2005).

In addition to sediment's effect on turbidity, trapped sediment homogenizes benthic habitat by infilling coarse materials through siltation. Siltation has a negative effect on the density and functional composition of benthic invertebrates and fish (Berkman and Rabeni, 1987; Rabení et al., 2005). Floodplain infilling by trapped sediment decreases depth, and often eliminates floodplain lakes altogether (Erskine et al., 1982). The loss of floodplain lakes markedly restructures fish assemblages due to the importance of this habitat. An additional manner by which sediment can decrease water quality is as a carrier of toxic substances, biodegradable organic matter, and nutrients (Ongley, 1996).

Eutrophication caused by increased nutrient loads has become a large problem, and is mainly attributable to agricultural practices and municipal sewage. Combined sewer overflows (CSO) serve 772 cities and over 40 million people in the United States (Aronson et al., 2012). Nitrogen and phosphorus are considered the primary limiting nutrients in freshwater systems; thus, have the most profound effect on eutrophication (Elser et al., 1990; Tilman, 1999). In the Mississippi River basin agriculture is responsible for approximately 70% of aquatic nitrogen and phosphorus inputs (Alexander et al., 2007). Increased nutrient levels restructure riverine communities from the bottom up and reduce overall biodiversity (Mainstone and Parr, 2002), particularly through its effect on dissolved oxygen (DO) (Bunn and Arthington, 2002; Hilton et al., 2006).

Hypoxia affects fishes in numerous ways. Direct mortality of fish and invertebrates results from extremely low DO (Arend et al., 2011; Wannamaker and Rice, 2000). Indirect effects on fish include changes in fish distribution, behavior, feeding rates, competitive interactions, and predation vulnerability (Wannamaker and Rice, 2000). Non-lethal hypoxia reduces growth rates and fecundity directly by lowering metabolic activity, and indirectly through energy lost on increased ventilation and movement rates (Kramer, 1987). Further affecting reproductive success, fish eggs and larvae generally have higher oxygen demands than adults have and are particularly vulnerable to hypoxia. Hypoxia can cause mortality and deformities to eggs and larvae at concentrations adults may tolerate (Davis, 1975).

There are numerous toxic pollutants affecting fish directly and indirectly through trophic cascades. Toxic chemicals include ammonia, pesticides, herbicides, various metals, phenols, solvents, chlorine, and pharmaceuticals. These toxins have been shown to reduce immunological function , increase susceptibility to parasitism, reduce growth rates, increase respiration rates, affect behavior, adversely affect reproduction, reduce swimming efficiency, and can result in direct mortality (Poulin, 1992; Randall and Tsui, 2002; Sprague, 1971; Zillich, 1972). Due to differential resistance to toxins among species, even sublethal concentrations have the potential to greatly restructure fish assemblages.

Habitat degradation and flow alterations are two other severe anthropogenic stressors affecting fish assemblages (Gordon et al., 2010; Pringle et al., 2000). Factors that alter the flow regime of large rivers include damming, channelization, and ground water withdrawal. These changes affect the magnitude, frequency, duration, and timing of flow. The alteration of the natural flow regime often leads to habitat loss, barriers to dispersal, changes in water temperature, changes in sediment flux, and ultimately to loss of fish biodiversity (Bunn and Arthington, 2002; Poff and Hart, 2002; Poff et al., 2007).

Dams are often constructed on large rivers to facilitate barge transport, generate electricity, and store water. During the 20th century over 45,000 dams greater than 15m high were constructed worldwide, and it is estimated there is one dam every 48km in third through seventh order river channels in the United States (Poff et al., 2007). Dams present physical obstructions to longitudinal fish movement. A recent review of the effects of dams on global fish biodiversity by Liermann and others (2012) found fish diversity on all continents, with the exception of Antarctica, to be at risk due to damming. Dams also cause the accumulation of fine sediments and toxic substances (Poff and Hart, 2002). The negative affect of dams can be seen in the immediate increase in fish and macrophyte biodiversity following removal (Bednarek, 2001).

Channelization also heavily contributes to habitat loss and alters biogeochemical processes. It is a broad group of engineered modifications which are used to control flooding, drain wetlands, alter channels for navigation, control bank erosion, and straighten channels (Brooker, 1985). Large rivers are commonly leveed for flood control and to aid in wetland draining for agriculture. Since European settlement 80-87% of the roughly 50 million hectares of wetlands drained in the contiguous United States have been drained for agricultural purposes (McCorvie and Lant, 1993). Channelization effects fish habitat by increasing overall discharge and sediment loads, and altering sediment depositional zones (Nakamura et al., 1997).

The exploitation of groundwater can also have ecosystem level consequences in large rivers. Groundwater is responsible for irrigating approximately 115 million hectares globally, and use is increasing (Siebert et al., 2010). The United States in 2005 withdrew 33% (14,600,000 gallons/day) of the public water supply, and 42% (53,500,000 gallons/day) of water used for irrigation is from groundwater (Kenny et al., 2009). Groundwater is directly responsible for surface water baseflow. Baseflow reductions can lead to higher concentrations of contaminants, and reduce habitat available to fish (Sophocleous, 2002).

Flow alteration and channelization can also strongly affect floodplain lakes, upon which many fishes are dependent for feeding areas, nurseries, and overwintering habitat (Power et al., 1996). When fishes are cut off from the floodplain there are marked decreases in biodiversity, and changes to assemblage structure (Agostinho et al., 2004; Jurajda, 1995). Dams reduce the flood pulse, levees restrict lateral connectivity of a river and its' floodplain, and groundwater withdrawals disrupt hyporheic exchange between a river and groundwater (Galat et al., 1998; Junk et al., 1989; Winter et al., 1998). Combined, these anthropogenic alterations reduce water quality and are capable of completely destroying floodplain lakes (Changming et al., 2001; Marshall, 2009; Pereira et al., 2009).

Many natural environmental factors can interact with human induced disturbances in the process of fish assemblage degradation or recovery. For purposes of this study, I focused on interannual climate variability. Climate variation plays a large role in fish assemblage structure, as regional climate variation is a key natural driver of biodiversity (Poff et al., 2007). Precipitation influences most aspects of the lotic environment, including sediment input, suspended sediment concentration, the concentrations of toxins and nutrients, velocity and discharge, and affects connectivity among main channel, floodplain lake, and terrestrial systems (Childers et al., 2006; Douglas, 1990; Neal et al., 1998; Ostrand and Wilde, 2001). The relationship between pollutant concentrations and precipitation is complex, as they are carried into the river through runoff, but rain can also dilute pollutant concentrations. Air temperature is the main determinant of river temperature, though human modifications can affect this relationship (Cairns Jr, 1972; Mohseni and Stefan, 1999; Poole and Berman, 2001; Webb et al., 2003). An analysis of historical records indicates rivers and streams throughout the U.S. are warming, particularly around urban areas (Kaushal et al., 2010). Temperature changes have the ability to restructure fish assemblages, as they effect DO concentrations, primary productivity, and fish metabolic rates (Carpenter and Dunham, 1985; Clarke and Johnston, 1999; Rahn, 1966).

Because of fish sensitivity to human disturbance, fish assemblages have long been used to assess river health (Doudoroff and Warren, 1957). Fish are widely distributed, often abundant, and fish sampling methods are well established. As a result, fish-based biological indicators, such as the Index of Biotic Integrity (IBI), have been most widely used to assess river ecosystem health (Roset et al., 2007).

Degradation and Recovery of Large Rivers

Humanity's ability to modify their environment is one of the greatest evolutionary adaptations in the biological world. This ability has led to substantial impacts on large river ecosystems. Absent anthropogenic alterations, fish diversity was primarily controlled by climate driven natural disturbance regimes (Poff et al., 2007). Meybeck (2003) posits that the natural controls have now been overtaken by a complex of “Anthropocene Syndromes”. In large river systems, combinations of pollution inputs, hydrological modifications, floodplain draining, and harvests of bottomland forests are primary causes of these syndromes. All of these modifications result in severe fish biodiversity loss, and consequentially impair the ability of rivers maintain ecosystem services (Bunn and Arthington, 2002; Nelson et al., 2009). A recent review found approximately 80% of the world’s human population is facing significant threats to water security, and this is coincident with 10,000–20,000 freshwater species extinct or at risk as a result of human impacts (Vörösmarty et al., 2010).

Developing countries are particularly at risk due to rapid human expansion with minimal investment in maintaining biodiversity (Vörösmarty et al., 2010). These countries are going through a similar process to the one that faced developed countries during industrialization. They are caught between economic growth and environmental preservation and unable to make the significant investments necessary to improve water quality and sustain riverine ecosystem integrity. Declining fish diversity during the 20th century was documented most developed nations, but success stories of increased water quality preceding successful recolonization have been documented (Malmqvist and Rundle, 2002). Two such success stories are the Rhine and Illinois rivers.

Rhine

The Rhine is a large European river, which drains areas of Switzerland, Germany, Austria, France, Belgium, Luxemburg, and The Netherlands. Hydrological modifications to the Rhine include dams, diversion canals, and channelization for boat traffic and hydropower (Plum and Schulte-Wülwer-Leidig, 2012; Raat, 2001). By 1970 the river was highly polluted with metals, nutrients, and sediment (Raat, 2001). The combination of pollution and habitat degradation led to the extirpation of many economically prized anadromous fish species, and restructured the fish assemblage by favoring pollution tolerant and invasive species (Raat, 2001).

Several international rehabilitation projects began in the 1980's aimed at increasing water quality, restoring habitat connectivity, and establishing self-sustaining populations of extirpated native species (Plum and Schulte-Wülwer-Leidig, 2012). Specific water quality parameters which have significantly improved include dissolved oxygen, concentrations of several heavy metals, and, as an indicator of decreased nutrient loading, chlorophyll a (Bij de Vaate et al., 2006). A basin survey conducted from 2006-2007 found that all known native fish species had returned with the exception of the Atlantic sturgeon (Plum and Schulte-Wülwer-Leidig, 2012). Through international cooperation, the series of projects were deemed successful, meeting many projected outcomes years in advance. Additionally, these series of successes helped serve as a template for the European Framework Directive which sets water quality and biodiversity goals for European Union member nations (Lepper, 2005; Plum and Schulte-Wülwer-Leidig, 2012).

Illinois River

Fish assemblages in the Illinois River Waterway (IRW) have been affected by all of the factors discussed. Its hydrology has been altered through installation of eight locks and dams, a canal creating an artificial connection with Lake Michigan, conversion of 40% of the floodplain to farming through leveeing, and maintenance of a 2.74 meter canal for navigation (Lian et al., 2012). Water quality in the IRW has been negatively affected by industrial activity, loss of the floodplain, and CSOs.

Hydrological modifications have affected the timing and duration of flood events in the IRW, and fish assemblage structure (Koel and Sparks, 2002; Lian et al., 2012). Sedimentation rates were increasing during the 1960's and 1970's (Steffeck et al., 1980). Combined, these two factors are cited as the cause of the collapse of macrophyte communities from the remaining floodplain lakes (Moore et al., 2010; Sparks et al., 1990). Toxic sediment in the IRW led to reductions in abundance and extirpations of numerous invertebrates, and affected fishes through bottom up trophic cascades (Sparks and Sandusky, 1983; Sparks et al., 1981). In the early 1970's ammonia and cyanide levels were sufficient to cause probable fish stress (Lubinski et al., 1974). Hypoxia and pollution led to near complete dominance of non-native common carp and goldfish in much of the IRW throughout the 1960's (Sparks and Lerczak, 1993). Climate has also been found to affect the IRW fish. Drought in the 1976-1977 water year raised toxin concentrations significantly, and flooding events altered assemblage structure through differential recruitment success (Koel and Sparks, 2002; Sparks et al., 1990).

The summarized negative impacts took a toll economically as well, leading to the collapse of a very successful commercial fishery (Sparks and Sandusky, 1981). In recent years, the IRW has shown significant improvement. Fish species richness, number of native species, and number of predatory species have all been increasing (Koel and Sparks, 2002; Lerczak et al., 1994; McClelland et al., 2012). Previous researchers have hypothesized the improvements in fish biodiversity were caused by improved water quality, likely due to the Clean Water Act (Pegg and McClelland, 2004).

Illinois River Long-term Monitoring and Research Needs

Studies analyzing multi-year fish assemblage alterations in large rivers are limited. The majority of large river research is confined to either spatial or short-term variation. Previous studies found long-term investigations critical for associating water quality changes with a biotic response (Elliott, 1990). Long-term monitoring allows discovery of delayed responses, detection of rare but significant event effects (e.g. floods, droughts), picks up alterations due to subtle long-term trends (e.g. improved water quality, climate change), helps prevent erroneous conclusions based on insufficient temporal data, and detects ecological processes which occur over long time scales (e.g. succession, invasion, population cycles) (Dodds et al., 2012; Elliott, 1990). Despite the advantages of long-term ecological monitoring, such programs are extremely rare in general and are particularly rare for large rivers (Strayer et al., 1986).

The Long-Term Illinois River Fish Population Monitoring Program, also called Long-term Electrofishing (LTEF), provides a structured and standardized data set to analyze temporal trends in large river fish assemblages. The LTEF study has sampled 26 fixed sites distributed throughout the IRW on a nearly continuous annual basis since 1957 (Koel and Sparks, 2002). Sampling utilizes alternating current (AC) boat electrofishing in August and September, and adheres to temperature water level restrictions (Koel and Sparks, 2002). The late summer sampling period permits young of year to reach a catchable size, while the temperature and water level requirements reduce the probability of fish dispersion and migration into overwintering habitat. These fixed methodologies reduce sampling bias, and permit powerful statistical analyses.

Previous analyses of LTEF data found that native fish, predatory fish, and species richness have been increasing since the mid 1970's throughout much of the IRW (Lerczak et al., 1994; McClelland et al., 2006; McClelland et al., 2012). These studies further showed spatial differences in the species composition among sections of the river. The major hydrological alterations to the IRW were completed over two decades before LTEF sampling began, with the last dams, Peoria and LaGrange, completed in 1939. This has led researchers to hypothesize the observed changes in LTEF fish assemblage data were due to factors other than hydrology, and pointed to water quality as being the driving factor behind these changes (Pegg and McClelland, 2004; Sparks and Sandusky, 1981).

Sparks and Lerczak (1993) found that the ammonia levels in the upper IRW were decreasing likely due to improvements in Chicago area sewage treatment, but Pegg and McClelland (2004) provide one of the only attempts to directly link the fish assemblage changes with water quality. They noted the lack of complementary water quality, and used sparse intermittent data obtained from the Environmental Protection Agency (EPA). Due to the paucity of the water quality data used by Pegg and McClelland (2004), statistical relationships were difficult to determine.

For this study, I will attempt to determine the effects of water quality on different biodiversity aspects of the IRW fish assemblages; including, species richness, predator abundance, native species abundance, and functional diversity. I will evaluate the strength of these relationships for different reaches, and ascertain the specific aspects of water quality influencing fish biodiversity. To address these research needs, I use complimentary water quality data previously uninvestigated for relationships with LTEF fish assemblage data. I also use regional climate data to account for some non-directional interannual variation in fish assemblages. In Chapter 2, I investigate relationships between these environmental factors (water quality and climate) and three aspects of fish assemblage structure; namely, the proportions of native species, the proportions of predator species, and species richness. In Chapter 3, I investigate the relationship between the environmental variables and functional diversity of the fish assemblages. In Chapter 4, I provide a general synthesis of this study and identify further research needs.

References

Agostinho, AA, Gomes, LC, Veríssimo, S and Okada, EK. 2004. Flood regime, dam regulation and fish in the Upper Paraná River: effects on assemblage attributes, reproduction and recruitment. *Reviews in Fish Biology and Fisheries* **14**: 11-19.

Alexander, RB, Smith, RA, Schwarz, GE, Boyer, EW, Nolan, JV and Brakebill, JW. 2007. Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River Basin. *Environmental Science & Technology* **42**: 822-830.

Allan, JD, Abell, R, Hogan, Z, Revenga, C, Taylor, BW, Welcomme, RL and Winemiller, K. 2005. Overfishing of inland waters. *BioScience* **55**: 1041-1051.

Arend, KK, Beletsky, D, DePinto, JV, Ludsin, SA, Roberts, JJ, Rucinski, DK, Scavia, D, Schwab, DJ and Höök, TO. 2011. Seasonal and interannual effects of hypoxia on fish habitat quality in central Lake Erie. *Freshwater Biology* **56**: 366-383.

Aronson, D, Weeks, J, Meylan, B, Guiney, PD and Howard, PH. 2012. Environmental release, environmental concentrations, and ecological risk of N, N-diethyl-m-toluamide (DEET). *Integrated Environmental Assessment and Management* **8**: 135-166.

Avitsur, S. 1960. On the history of the exploitation of water power in Eretz-Israel. *Israel Exploration Journal* **10**: 37-45.

Banerjee, L. 2010. Effects of flood on agricultural productivity in Bangladesh. *Oxford Development Studies* **38**: 339-356.

Bayley, PB. 1995. Understanding large river: floodplain ecosystems. *BioScience* **45**: 153-158.

Bednarek, AT. 2001. Undamming rivers: a review of the ecological impacts of dam removal. *Environmental Management* **27**: 803-814.

Berkman, HE and Rabeni, CF. 1987. Effect of siltation on stream fish communities. *Environmental Biology of Fishes* **18**: 285-294.

Bij de Vaate, A, Breukel, R and Van der Velde, G. 2006. Long-term developments in ecological rehabilitation of the main distributaries in the Rhine delta: fish and macroinvertebrates. In: *Living Rivers: Trends and Challenges in Science and Management*. Springer, pp 229-242

Bockstael, NE, Freeman, AM, Kopp, RJ, Portney, PR and Smith, VK. 2000. On measuring economic values for nature. *Environmental Science & Technology* **34**: 1384-1389.

Breukelaar, AW, Lammens, EH, Breteler, JGK and Tatrai, I. 1994. Effects of benthivorous bream (*Abramis brama*) and carp (*Cyprinus carpio*) on sediment resuspension and concentrations of nutrients and chlorophyll a. *Freshwater Biology* **32**: 113-121.

Brooker, MP. 1985. The ecological effects of channelization. *The Geographical Journal* **151**: 63-69.

Brown, GM and Shogren, JF. 1998. Economics of the Endangered Species Act. *The Journal of Economic Perspectives* **12**: 3-20.

Brusven, M and Rose, S. 1981. Influence of substrate composition and suspended sediment on insect predation by the torrent sculpin, *Cottus rhotheus*. *Canadian Journal of Fisheries and Aquatic Sciences* **38**: 1444-1448.

Bunn, SE and Arthington, AH. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management* **30**: 492-507.

Cairns Jr, J. 1972. Coping with heated waste water discharges from steam-electric power plants. *BioScience*: 411-423.

Cardinale, BJ, Srivastava, DS, Emmett Duffy, J, Wright, JP, Downing, AL, Sankaran, M and Jouseau, C. 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* **443**: 989-992.

Carpenter, EJ and Dunham, S. 1985. Nitrogenous nutrient uptake, primary production, and species composition of phytoplankton in the Carmans River estuary, Long Island, New York. *Limnology and Oceanography*: 513-526.

Changming, L, Jingjie, Y and Kendy, E. 2001. Groundwater exploitation and its impact on the environment in the North China Plain. *Water International* **26**: 265-272.

Chapin III, FS, Walker, BH, Hobbs, RJ, Hooper, DU, Lawton, JH, Sala, OE and Tilman, D. 1997. Biotic control over the functioning of ecosystems. *Science* **277**: 500-504.

Childers, DL, Boyer, JN, Davis, SE, Madden, CJ, Rudnick, DT and Sklar, FH. 2006. Relating precipitation and water management to nutrient concentrations in the oligotrophic "upside-down" estuaries of the Florida Everglades. *Limnology and Oceanography* **51**: 602-616.

Church, RL and Bell, TL. 1988. An Analysis of Ancient Egyptian Settlement Patterns Using Location–Allocation Covering Models. *Annals of the Association of American Geographers* **78**: 701-714.

Clarke, A and Johnston, NM. 1999. Scaling of metabolic rate with body mass and temperature in teleost fish. *Journal of Animal Ecology* **68**: 893-905.

Davis, JC. 1975. Minimal dissolved oxygen requirements of aquatic life with emphasis on Canadian species: a review. *Journal of the Fisheries Board of Canada* **32**: 2295-2332.

Dodds, WK, Robinson, CT, Gaiser, EE, Hansen, GJ, Powell, H, Smith, JM, Morse, NB, Johnson, SL, Gregory, SV and Bell, T. 2012. Surprises and insights from long-term aquatic data sets and experiments. *BioScience* **62**: 709-721.

Doudoroff, P and Warren, CE. 1957. Biological indices of water pollution, with special reference to fish populations. In: *Biological Problems in Water Pollution*. US Public Health Service, Cincinnati, OH

Douglas, I. 1990. Sediment transfer and siltation. In: The Earth as Transformed by Human Action. Press Syndicate of the University of Cambridge, New York, NY, pp 215-234

Elliott, J. 1990. The need for long-term investigations in ecology and the contribution of the Freshwater Biological Association. *Freshwater Biology* **23**: 1-5.

Elser, JJ, Marzolf, ER and Goldman, CR. 1990. Phosphorus and nitrogen limitation of phytoplankton growth in the freshwaters of North America: a review and critique of experimental enrichments. *Canadian Journal of Fisheries and Aquatic Sciences* **47**: 1468-1477.

Erskine, W, Melville, M, Page, K and Mowbray, P. 1982. Cutoff and oxbow lake. *The Australian Geographer* **15**: 174-180.

Forbes, SA. 1925. The lake as a microcosm. *Biology of Water Pollution* **15**: 3.

Galat, DL, Fredrickson, LH, Humburg, DD, Bataille, KJ, Bodie, JR, Dohrenwend, J, Gelwicks, GT, Havel, JE, Helmers, DL and Hooker, JB. 1998. Flooding to restore connectivity of regulated, large-river wetlands. *BioScience* **48**: 721-733.

Gelwick, FP, Stock, MS and Matthews, WJ. 1997. Effects of fish, water depth, and predation risk on patch dynamics in a North-Temperate river ecosystem. *Oikos* **80**: 382-398.

Gonzales, D, Searcy, EM and Ekşioğlu, SD. 2013. Cost analysis for high-volume and long-haul transportation of densified biomass feedstock. *Transportation Research Part A: Policy and Practice* **49**: 48-61.

Gordon, LJ, Finlayson, CM and Falkenmark, M. 2010. Managing water in agriculture for food production and other ecosystem services. *Agricultural Water Management* **97**: 512-519.

Harris, M. 2013. Philadelphia's water supply: how human capital and collaboration can overcome significant man-made challenges. *Graduate Studies Journal of Organizational Dynamics* **2**: 2.

Hilton, J, O'Hare, M, Bowes, MJ and Jones, JI. 2006. How green is my river? A new paradigm of eutrophication in rivers. *Science of the Total Environment* **365**: 66-83.

Holmlund, CM and Hammer, M. 1999. Ecosystem services generated by fish populations. *Ecological Economics* **29**: 253-268.

Johnson, JE and Hines, RT. 1999. Effect of suspended sediment on vulnerability of young razorback suckers to predation. *Transactions of the American Fisheries Society* **128**: 648-655.

Junk, WJ, Bayley, PB and Sparks, RE. 1989. The flood pulse concept in river-floodplain systems. *Canadian Special Publication of Fisheries and Aquatic Sciences* **106**: 110-127.

Jurajda, P. 1995. Effect of channelization and regulation on fish recruitment in a flood plain river. *Regulated Rivers: Research & Management* **10**: 207-215.

Kaushal, SS, Likens, GE, Jaworski, NA, Pace, ML, Sides, AM, Seekell, D, Belt, KT, Secor, DH and Wingate, RL. 2010. Rising stream and river temperatures in the United States. *Frontiers in Ecology and the Environment* **8**: 461-466.

Kenny, JF, Barber, NL, Hutson, SS, Linsey, KS, Lovelace, JK and Maupin, MA. 2009. Estimated use of water in the United States in 2005. US Geological Survey Reston, VA.

Koel, TM and Sparks, RE. 2002. Historical patterns of river stage and fish communities as criteria for operations of dams on the Illinois River. *River Research and Applications* **18**: 3-19.

Kozicky, EL. 1977. Tomorrow's hunters: gadgeteers or sportsmen? *Wildlife Society Bulletin* **5**: 175-178.

Kramer, DL. 1987. Dissolved oxygen and fish behavior. *Environmental Biology of Fishes* **18**: 81-92.

Kratzer, EB, Jackson, JK, Arscott, DB, Aufdenkampe, AK, Dow, CL, Kaplan, L, Newbold, J and Sweeney, BW. 2006. Macroinvertebrate distribution in relation to land use and water chemistry in New York City drinking-water-supply watersheds. *Journal of the North American Benthological Society* **25**: 954-976.

Kummu, M, de Moel, H, Ward, PJ and Varis, O. 2011. How close do we live to water? A global analysis of population distance to freshwater bodies. *PLoS ONE* **6**: e20578.

Kwak, TJ. 1988. Lateral movement and use of floodplain habitat by fishes of the Kankakee River, Illinois. *American Midland Naturalist*: 241-249.

Lepper, P. 2005. Manual on the methodological framework to derive environmental quality standards for priority substances in accordance with Article 16 of the Water Framework Directive (2000/60/EC). Fraunhofer-Institute Molecular Biology and Applied Ecology, Schmallenberg, Germany.

Lerczak, TV, Sparks, RE and Blodgett, DK. 1994. The long-term Illinois River fish population monitoring program project F-101-R. Illinois Natural History Survey, Champaign, IL.

Lian, Y, You, J, Sparks, R and Demissie, M. 2012. Impact of human activities to hydrologic alterations on the Illinois River. *Journal of Hydrologic Engineering* **17**: 537-546.

Liang, W, Wu, Z-b, Cheng, S-p, Zhou, Q-h and Hu, H-y. 2003. Roles of substrate microorganisms and urease activities in wastewater purification in a constructed wetland system. *Ecological Engineering* **21**: 191-195.

Liermann, CR, Nilsson, C, Robertson, J and Ng, RY. 2012. Implications of dam obstruction for global freshwater fish diversity. *BioScience* **62**: 539-548.

Lubinski, KS, Sparks, RE and Jahn, LA. 1974. The development of toxicity indices for assessing the quality of the Illinois River. Illinois Natural History Survey, Champaign, IL.

Mainstone, CP and Parr, W. 2002. Phosphorus in rivers: ecology and management. *Science of the Total Environment* **282**: 25-47.

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

Marshall, DW. 2009. Surveys of Lower Wisconsin River floodplain lakes: Sauk county, Wisconsin. *Lakes Planning Grant Study*: 1-36.

Mason, R, Laporte, J-M and Andres, S. 2000. Factors controlling the bioaccumulation of mercury, methylmercury, arsenic, selenium, and cadmium by freshwater invertebrates and fish. *Archives of Environmental Contamination and Toxicology* **38**: 283-297.

McClelland, MA, Pegg, MA and Spier, TW. 2006. Longitudinal patterns of the Illinois Waterway fish community. *Journal of Freshwater Ecology* **21**: 91-99.

McClelland, MA, Sass, GG, Cook, TR, Irons, KS, Michaels, NN, O'Hara, TM and Smith, CS. 2012. The long-term Illinois River fish population monitoring program. *Fisheries* **37**: 340-350.

McCorvie, MR and Lant, CL. 1993. Drainage district formation and the loss of Midwestern wetlands, 1850-1930. *Agricultural History* **67**: 13-39.

Meybeck, M. 2003. Global analysis of river systems: from Earth system controls to Anthropocene syndromes. *Philosophical Transactions of the Royal Society of London: Biological Sciences* **358**: 1935-1955.

Meybeck, M. 2004. The global change of continental aquatic systems: dominant impacts of human activities. *Water Science and Technology* **49**: 73.

Mohseni, O and Stefan, H. 1999. Stream temperature/air temperature relationship: a physical interpretation. *Journal of Hydrology* **218**: 128-141.

Moigne, GL and Priscoli, JD. 1992. Water, rivers, and mankind: challenges for the coming millennium. *Water International* **17**: 81-87.

Moore, M, Romano, SP and Cook, T. 2010. Synthesis of upper Mississippi River system submersed and emergent aquatic vegetation: past, present, and future. *Hydrobiologia* **640**: 103-114.

Nakamura, F, Sudo, T, Kameyama, S and Jitsu, M. 1997. Influences of channelization on discharge of suspended sediment and wetland vegetation in Kushiro Marsh, northern Japan. *Geomorphology* **18**: 279-289.

Neal, C, Harrow, M and Williams, RJ. 1998. Dissolved carbon dioxide and oxygen in the River Thames: spring–summer 1997. *Science of the Total Environment* **210**: 205-217.

Nelson, E, Mendoza, G, Regetz, J, Polasky, S, Tallis, H, Cameron, D, Chan, KMA, Daily, GC, Goldstein, J, Kareiva, PM, Lonsdorf, E, Naidoo, R, Ricketts, TH and Shaw, M. 2009. Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment* **7**: 4-11.

O'Connor, GC. 2013. America's quest for clean energy: is hydropower relevant? *The Electricity Journal* **26**: 15-24.

Ongley, ED. 1996. Control of water pollution from agriculture. *FaO irrigation and drainage Paper* **55**: 1-101.

Ostrand, KG and Wilde, GR. 2001. Temperature, dissolved oxygen, and salinity tolerances of five prairie stream fishes and their role in explaining fish assemblage patterns. *Transactions of the American Fisheries Society* **130**: 742-749.

Pegg, MA and McClelland, MA. 2004. Spatial and temporal patterns in fish communities along the Illinois River. *Ecology of Freshwater Fish* **13**: 125-135.

Pereira, LS, Cordery, I and Iacovides, I. 2009. Groundwater use and recharge. In: *Coping with Water Scarcity*. Springer, pp 133-174

Petes, LE, Brown, AJ and Knight, CR. 2012. Impacts of upstream drought and water withdrawals on the health and survival of downstream estuarine oyster populations. *Ecology and Evolution* **2**: 1712-1724.

Plum, N and Schulte-Wülwer-Leidig, A. 2012. From a sewer into a living river: the Rhine between Sandoz and Salmon. *Hydrobiologia*: 1-12.

Poff, NL and Hart, DD. 2002. How dams vary and why it matters for the emerging science of dam removal: an ecological classification of dams is needed to characterize how the tremendous variation in the size, operational mode, age, and number of dams in a river basin influences the potential for restoring regulated rivers via dam removal. *BioScience* **52**: 659-668.

Poff, NL, Olden, JD, Merritt, DM and Pepin, DM. 2007. Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences* **104**: 5732-5737.

Poole, GC and Berman, CH. 2001. An ecological perspective on in-stream temperature: natural heat dynamics and mechanisms of human-caused thermal degradation. *Environmental Management* **27**: 787-802.

Poulin, R. 1992. Toxic pollution and parasitism in freshwater fish. *Parasitology Today* **8**: 58-61.

Power, ME, Dietrich, WE and Finlay, JC. 1996. Dams and downstream aquatic biodiversity: potential food web consequences of hydrologic and geomorphic change. *Environmental Management* **20**: 887-895.

Pringle, CM, Freeman, MC and Freeman, BJ. 2000. Regional effects of hydrologic alterations on riverine macrobiota in the New World: tropical-temperate comparison. *BioScience* **50**: 807-823.

Raat, AJ. 2001. Ecological rehabilitation of the Dutch part of the River Rhine with special attention to the fish. *Regulated Rivers: Research & Management* **17**: 131-144.

Rabení, CF, Doisy, KE and Zweig, LD. 2005. Stream invertebrate community functional responses to deposited sediment. *Aquatic Sciences* **67**: 395-402.

Rahn, H. 1966. Aquatic gas exchange: theory. *Respiration physiology* **1**: 1-12.

Randall, D and Tsui, T. 2002. Ammonia toxicity in fish. *Marine Pollution Bulletin* **45**: 17-23.

Richards, WH, Koeck, R, Gersonde, R, Kuschnig, G, Fleck, W and Hochbichler, E. 2012. Landscape-Scale forest management in the municipal watersheds of Vienna, Austria, and Seattle, USA: commonalities despite disparate ecology and history. *Natural Areas Journal* **32**: 199-207.

Roset, N, Grenouillet, G, Goffaux, D, Pont, D and Kestemont, P. 2007. A review of existing fish assemblage indicators and methodologies. *Fisheries Management and Ecology* **14**: 393-405.

Ross, ST and Baker, JA. 1983. The response of fishes to periodic spring floods in a Southeastern stream. *American Midland Naturalist*: 1-14.

Saravanan, V, Mollinga, PP and Bogardi, JJ. 2011. Global change, wastewater and health in fast growing economies. *Current Opinion in Environmental Sustainability* **3**: 461-466.

Schramm Jr, HL. 2004. Status and management of Mississippi River fisheries. *Proceedings of the Second International Symposium on the Management of Large Rivers for Fisheries. Food and Agriculture Organization of the United Nations.* Bangkok, Thailand.

Schroeder, S. 1999. Maize productivity in the Eastern woodlands and Great Plains of North America. *American Antiquity* **64**: 499-516.

Shingles, A, McKenzie, D, Claireaux, G and Domenici, P. 2005. Reflex cardioventilatory responses to hypoxia in the flathead gray mullet (*Mugil cephalus*) and their behavioral modulation by perceived threat of predation and water turbidity. *Physiological and Biochemical Zoology* **78**: 744-755.

Siebert, S, Burke, J, Faures, J-M, Frenken, K, Hoogeveen, J, Döll, P and Portmann, FT. 2010. Groundwater use for irrigation: a global inventory. *Hydrology and Earth System Sciences* **14**: 1863-1880.

Smakhtin, V, Revenga, C and Döll, P. 2004. A pilot global assessment of environmental water requirements and scarcity. *Water International* **29**: 307-317.

Sophocleous, M. 2002. Interactions between groundwater and surface water: the state of the science. *Hydrogeology Journal* **10**: 52-67.

Sparks, RE. 1995. Need for ecosystem management of large rivers and their floodplains. *BioScience* **45**: 168-182.

Sparks, RE, Bayley, PB, Kohler, SL and Osborne, LL. 1990. Disturbance and recovery of large floodplain rivers. *Environmental Management* **14**: 699-709.

Sparks, RE and Lerczak, TV. 1993. Recent trends in the Illinois River indicated by fish populations. Illinois Natural History Survey, Champaign, IL.

Sparks, RE and Sandusky, MJ. 1981. Identification of factors responsible for decreased production of fish food organisms in the Illinois and Mississippi Rivers. Illinois Natural History Survey, Champaign, IL.

Sparks, RE and Sandusky, MJ. 1983. Identification of the water quality factors which prevent fingernail clams from recolonizing the Illinois River phase: phase III. Illinois Natural History Survey, Champaign, IL.

Sparks, RE, Sandusky, MJ and Paparo, AA. 1981. Identification of the water quality factors which prevent fingernail clams from recolonizing the Illinois River: phase I. Illinois Natural History Survey, Champaign, IL.

Sprague, J. 1971. Measurement of pollutant toxicity to fish—III: sublethal effects and “safe” concentrations. *Water Research* **5**: 245-266.

Srivastava, DS and Vellend, M. 2005. Biodiversity-ecosystem function research: is it relevant to conservation? *Annual Review of Ecology, Evolution, and Systematics* **36**: 267-294.

Steffeck, DW, Paveglio, FL, Bellrose, FC and Sparks, RE. 1980. Effects of decreasing water depths on the sedimentation rate of illinois river bottomland lakes. *Journal of the American Water Resources Association* **16**: 553-555.

Strayer, D, Glitzenstein, JS, Jones, CG, Kolasa, J, Likens, GE, McDonnell, MJ, Parker, GG, Pickett, ST and Arboretum, MFC. 1986. Long-term ecological studies: an illustrated account of their design, operation, and importance to ecology. Institute of Ecosystem Studies, Millbrook, NY.

Tilman, D. 1999. Global environmental impacts of agricultural expansion: the need for sustainable and efficient practices. *Proceedings of the National Academy of Sciences* **96**: 5995-6000.

Urbansky, ET, Magnuson, ML, Kelty, CA and Brown, SK. 2000. Perchlorate uptake by salt cedar (*Tamarix ramosissima*) in the Las Vegas Wash riparian ecosystem. *Science of the Total Environment* **256**: 227-232.

US Fish and Wildlife Service. 2011. 2011 National survey of fishing, hunting, and wildlife-associated recreation. Washington, DC.

Van der Oost, R, Beyer, J and Vermeulen, NP. 2003. Fish bioaccumulation and biomarkers in environmental risk assessment: a review. *Environmental Toxicology and Pharmacology* **13**: 57-149.

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

Vörösmarty, CJ, McIntyre, P, Gessner, MO, Dudgeon, D, Prusevich, A, Green, P, Glidden, S, Bunn, SE, Sullivan, CA and Liermann, CR. 2010. Global threats to human water security and river biodiversity. *Nature* **467**: 555-561.

Vought, LBM, Dahl, J, Pedersen, CL and Lacoursière, JO. 1994. Nutrient retention in riparian ecotones. *Ambio* **23**: 342-348.

Vrtiska, MP, Gammonley, JH, Naylor, LW and Raedeke, AH. 2013. Economic and conservation ramifications from the decline of waterfowl hunters. *Wildlife Society Bulletin* **37.2**: 380-388.

Wannamaker, CM and Rice, JA. 2000. Effects of hypoxia on movements and behavior of selected estuarine organisms from the southeastern United States. *Journal of Experimental Marine Biology and Ecology* **249**: 145-163.

Webb, B, Clack, P and Walling, D. 2003. Water–air temperature relationships in a Devon river system and the role of flow. *Hydrological Processes* **17**: 3069-3084.

Winter, TC, Harvey, JW, Franke, OL and Alley, WM. 1998. *Ground water and surface water: a single resource*. U.S. Government Printing Office, Denver, CO.

Zillich, JA. 1972. Toxicity of combined chlorine residuals to freshwater fish. *Water Pollution Control Federation* **44**: 212-220.

Zimny, J, Michalak, P, Bielik, S and Szczotka, K. 2013. Directions in development of hydropower in the world, in Europe and Poland in the period 1995–2011. *Renewable and Sustainable Energy Reviews* **21**: 117-130.

Chapter 2

Long-term Changes of Fish Assemblages in Relation to Water Quality and Climate Variability in the Illinois River Waterway

Abstract

For over a hundred years, a range of human disturbances have affected the hydrology, habitat, and water quality of the Illinois River Waterway's (IRW) ecosystem. Soon after implementation of the Clean Water Act in the 1970s, substantial changes in species diversity and composition were observed. Fish assemblage alterations have generally been attributed to water quality improvements. However, these observations have not been statistically linked to specific water quality variables. We obtained comprehensive water quality data for the time period from 1983 to 2010 for the upper 207 km of the Upper Illinois River Waterway in order to assess linkages with multiple assemblage attributes. To account for considerable interannual non-directional variation in fish data, regional weather data related to precipitation and temperature over the same period were also compiled. We used random forests (RF) and multiple linear regressions (MLR) to examine the relationships of several fish-assemblage attributes with both water quality and climate variables. The assemblage attributes examined include species richness, predator-species biomass and abundance, and native-fish biomass and abundance. Model performances varied among reaches and response variables ($0 \leq R^2 \leq 0.73$ for RF; $0.14 \leq R^2 \leq 0.94$ for MLR). Both types of models indicated dissolved oxygen, water clarity, and nutrients were the main factors affecting assemblage structure, while climate plays a subordinate role. We conclude that pollution mitigation has played a substantial role in allowing the recovery of the IRW fish assemblages.

Introduction

Freshwater fishes account for approximately 40% of global fish species diversity and 25% of total vertebrate diversity even though freshwater habitats only account for 0.08% of the Earth's surface (Dudgeon et al., 2005). Fish diversity patterns reflect physical, chemical, and biological processes operating at multiple spatial and temporal scales. At broad scales, factors such as continental drift, glaciation, and climate history are critical (Levêque et al., 2008). At finer temporal and spatial scales, climate variability, hydrological regime, habitat diversity, trophic interactions, and water chemistry often shape fish diversity (Guégan et al., 1998; Oberdorff et al., 2011; Poff et al., 2007).

In modern times, human impacts have become an important force driving patterns of fish diversity. Degradation of the world's rivers occurred at an alarming rate following the industrial revolution (Donohoe, 2003; Douglas, 1990). The degradation is a symptom induced by human alterations, and has been described as a suite of "Anthropocene syndromes" caused by hydrologic modifications, pollution, and land use changes (Meybeck, 2003). These changes have impaired fish assemblages and their biological functions in most rivers (Boulton, 1999; Bunn and Arthington, 2002).

Many societies have recognized the adverse impacts they are having on river ecosystems. Accordingly, attitudes have shifted to reflect a greater concern for the future wellbeing of the rivers upon which they depend, and these desires have been expressed through legislation (e.g. US Clean Water Act, US Endangered Species Act, and EU Water Framework Directive). These changes in perception and accompanying legislation have begun to reverse the trend of biological degradations in many large rivers. Though certain necessary alterations (e.g. agriculture, transportation, industry) are likely to persist, reductions in water pollution are obtainable. Indeed, pollution mitigation has already shown significant and positive impacts to many large river ecosystems (Douglas and McCreanor, 1990; Langford et al., 2010; Romano et al., 1992).

Long-term monitoring is critical for detecting changes in fish assemblages. This is especially true for changes which are temporally directional, but slow or associated with substantial stochastic variation or infrequent events (e.g., severe droughts or floods) (Turner et al., 2003). However, these monitoring programs are costly and rare, especially for large rivers (Elliott, 1990; Strayer et al., 1986). The Long-term Illinois River Fish Population Monitoring Program (LTEF) initiated by the Illinois Natural History Survey (INHS) in 1957 to monitor fish populations in the Illinois River, Illinois USA, is one of the few such programs worldwide (McClelland et al., 2012). These data offer opportunities to investigate the effects of human disturbances on large river fish assemblages.

As is the case with the majority of large rivers across the world, the Illinois River has been heavily modified to meet human demands placed upon it. Alterations have impacted its hydrological regime, habitat complexity, watershed land cover, and water quality (Lian et al., 2012; Lubinski et al., 1974; Moore et al., 2010). During the 20th century, numerous alterations in the structure of the Illinois River Waterway (IRW) fish assemblages were noted as coincident with the impairments caused by the anthropogenic disturbances. Changes in the flow regime have been implicated in altering fish diversity (Koel and Sparks, 2002). Toxic sediment caused the extirpation of several invertebrate species leading to declines in fish condition, abundance, and played a role in the collapse of a once highly successful commercial fishery through trophic cascades (Sparks and Sandusky, 1981). As evidence of degradation, throughout the 1960's much of the IRW was extremely depauperate, dominated by pollution-tolerant common carp (*Cyprinus carpio*) and goldfish (*Carassius auratus*) (Sparks and Lerczak, 1993).

More recently, water quality improvements within the Upper IRW due to enhanced sewage treatment, and reduction of untreated combined sewage overflow (CSO) wastewater release through the Deep Tunnel Project have been reported (Sparks and Lerczak, 1993; Theriot and Tzoumis, 2005). Fish species richness, abundances of native species, and abundance of predator species have all been increasing (Koel and Sparks, 2002; Lerczak et al., 1994; Pegg and McClelland, 2004). Pegg and McClelland (2004) attempted to link changes in the fish assemblages with six water quality variables taken intermittently between 1958 and 1991. These authors failed to identify strong statistical relationships between diversity or relative abundances with water quality due, in part, to the paucity of the water quality data. Additionally, their assessment did not permit determination of the relative influence of specific water quality variables on the assemblages.

Against this backdrop of improved water quality and fish diversity, some unanswered questions remain. For example, how important is water quality in explaining observed fish assemblage changes relative to other factors such as interannual weather variation? In addition, how does the importance of water quality for fish assemblages vary from upper to downstream reaches? The answers to these questions are important to future management of the IRW, and to further ecological knowledge regarding the recovery of fish assemblages in heavily modified large river ecosystems. Properly addressing these questions requires detailed water quality and climate data to match the fish sampling data, and use of robust statistical modeling approaches which can handle complex fish-environmental relationships.

Our study examines the upper 207km of the IRW for the period 1983-2010. Water quality data were obtained from the Metropolitan Water Reclamation District of Greater Chicago (MWRD). MWRD has monitored water quality in the IRW mainstem from Chicago to Peoria using fixed sites and methodologies since 1983. In addition, we compiled climate data for modeling procedures from the National Oceanic and Atmospheric Administration (NOAA). We characterize statistical relationships between a suite of environmental variables composed of 12 water quality and 16 climate variables, and several fish assemblage attributes. Our specific questions include: 1) How did different assemblage attributes (e.g. species richness, proportional changes in predatory or native species) change over time?, 2) Which water quality variables best explained observed long-term changes in a given assemblage attribute?, 3) How did interannual weather variation contribute to fish assemblage changes?, 4) Does the relative importance of water quality differ among different reaches of the IRW?

Methods

Study Area: Illinois River Waterway

The Illinois River is a 439 km braided river located in Illinois, United States, with a drainage area of 72,701 km^2 formed by the confluence of the Des Plaines and Kankakee Rivers. After being joined by several other tributaries, the Illinois River drains into the

Mississippi River near Alton, Illinois. A range of human activities has heavily altered the river since the early nineteenth century. In the 1830s a canal was constructed, connecting the Illinois River and Lake Michigan basins for boat passage. This was followed by water diversion through the Chicago Sanitary and Ship Canal (CSSC) in 1900 to redirect the flow of wastewater from Lake Michigan, the primary source of potable water for the area, into the Illinois River system. To facilitate shipping, seven locks and dams were built to allow a 2.74 m navigation channel depth to be maintained from the Mississippi confluence to Lake Michigan in the 1930s. Several cities within the basin, including Chicago, also make use of combined sewage overflow (CSO) systems, which release untreated sewage when overwhelmed with precipitation. Furthermore, approximately 40% of the river valley, originally dominated by tall-grass prairie and wetlands, was converted to farm land through levee and drainage districts by 1915 (Lian et al., 2012). These modifications significantly altered the flow regime, and impaired water quality (Lian et al., 2012; Lubinski et al., 1974; Moore et al., 2010).

For purposes of this study, we examined the river's upper four reaches delineated by five lock and dam (L&D) complexes; Dresden (Brandon Road L&D to Dresden L&D), Marseilles (Dresden L&D to Marseilles L&D), Starved Rock (Marseilles L&D to Starved Rock L&D), and Peoria (Starved Rock L&D to Peoria L&D) (Figure 2.1). While fish assemblage data have been collected throughout the entire river, we limited analyses to these four reaches for which we were able to attain complimentary water quality data (1983-2010). This division of the river has been well established by previous studies (Koel and Sparks, 2002; McClelland et al., 2006; McClelland and Sass, 2012; McClelland et al., 2012; Pegg and McClelland, 2004; Sparks, 2010; Sparks and Lerczak, 1993). These studies demonstrated significant differences among reaches in terms of flow regime, elevation, slope, adjoining land cover, geology, chemistry, and fish diversity (Butts et al., 1975; Koel and Sparks, 2002; Lian et al., 2012; McClelland et al., 2006; Pegg and McClelland, 2004). The dams create a physical barrier to fish passage, nutrient transport, and create clear hydrological divisions. Moreover, the division helps overcome some technical difficulties. Fish survey sites were occasionally sampled unequally, and reaches contained different numbers of sites to compensate for differences in reach length (McClelland et al., 2006). Pooling of data into reaches allows comparisons among reaches despite this sampling inequality. Further, reach averaging more accurately represents the reach than single sample locations, which are more vulnerable to stochastic variations than averaged data. Lastly, environmental data were not available for each fish survey site. The reach level approach averages out environmental differences among reaches, permits comparison between reaches despite differing numbers of sampling sites within each, and accommodated inclusion of the available environmental data.

Data Sources

Fish Assemblage Diversity: Long Term Fish Population Monitoring Program

The Long Term Fish Population Monitoring Program (LTEF) has collected fish assemblage samples using alternating current electrofishing on a nearly continuous annual basis since 1957. Each year, electrofishing was conducted at fixed-site locations in the IRW. From upstream to downstream, Dresden Reach contained two sites, Marseilles Reach three sites, Starved Rock Reach two sites, and Peoria Reach eight sites. Sites were chosen initially for their accessibility and likelihood of containing high levels of fish diversity based upon habitat characteristics (Pegg and McClelland, 2004). Each site was sampled once a year within a six-week period from late August to early October during low water when water temperatures exceeded 14° C. The early fall sampling period allows young of year fish to attain a catchable size. The water level and temperature requirements reduce fish dispersion, by limiting access to off-channel habitat and reducing the probability of fish migration to overwintering sites.

We described fish assemblages with seven attributes; 1) species richness (S), 2) proportion of predator species (S_p), 3) proportion of predator abundance (A_p), 4) proportion of predator biomass (W_p), 5) proportion of native species (S_n), 6) proportion of native abundance (A_n), and 7) proportion of native biomass(W_n). The classifications of native and predator species were based upon characters identified in the FishTrait database (Frimpong and Angermeier, 2009). These authors categorized species as native or non-native using regional geographic distribution maps. FishTrait's trophic trait of fishes that prey upon larger fishes, crayfishes, crabs, and frogs (FSHCRCRB) was used to define fish as predatory.

Species richness is a measure of species diversity that has been the most commonly used for both conservation and assessment of river health (Galat and Zweimüller, 2001; Harris and Silveira, 1999). Changes in the proportion of native species is representative of environmental degradation due to the evolutionary history with which native fish species have adapted to fill niches created by the complex habitat in large rivers. Reductions in native fish species, or inversely increases in non-native species, is often an indication of habitat alterations and water quality degradation (Byers, 2002). This often leads to exploitation by tolerant non-native generalists (Galat and

Zweimüller, 2001). We examined predatory fish due to their importance as sport fishes and their ability to reflect ecosystem productivity as secondary consumers (Wootton and Power, 1993)

Abundance (A) and biomass (W) were used to calculate proportional assemblage changes separately, as they respond differently to the environment. Abundance is better able to detect reestablishment of small-bodied native fishes, but biomass more accurately reflects the density of recreationally important larger bodied native predatory species. Additionally, abundance is more susceptible to interannual variations than biomass, especially climate-induced variability (Harrell, 1978; Minns et al., 1996; Raibley et al., 1997). Biomass may be more reflective of increased water quality, as it is less prone to the stochastic variation caused by climate variability.

We made two modifications to the original LTEF data as to better serve our research objectives. First, small-bodied fish were often not weighed by LTEF. To obtain a better estimate of biomass, we constructed growth curves for each species using all available LTEF data. Then lengths were used to estimate weights of individuals. A minimum weight of one gram was assigned to all fish observations, the lowest unit of measurement used by LTEF. While this may overestimate the biomass of some individuals weighing less than a gram, it allows the more abundant smaller fishes to be represented in the models. Second, electrofishing duration was typically one hour, but exceptions occurred. To standardize comparisons among years, catch per unit effort (CPUE) was calculated both as biomass in grams per hour and abundance in number per hour for each species during each sampling event. CPUE data were then used to calculate proportions of native and predatory fish in terms of number of species, abundance, and biomass.

Before modeling the changes in the fish assemblage attributes, we assessed their differences among reaches and years using ANOVA and Pearson correlation (Crawley, 2012; Hay-Jahans, 2011). The base R package "stats" was used in R 2.14.1 software to calculate Pearson's r and carry out ANOVA tests.

Water Quality Parameters: Metropolitan Water Reclamation District

With the exception of 1988, since 1983 the MWRD has collected water quality data at 49 fixed-sites located along 214 km of the upper IRW. Water samples were collected several times annually from May to October. Collection sites were at generally shallow and relatively narrow channels with plug-flow characteristics assumed to be well mixed. At each site water temperature (°C), and dissolved oxygen (mg/l) were measured using a handheld sensor (YSI 6600). A single water grab sample was also taken in the center of the waterway a meter below the surface, and preserved for laboratory analysis of nutrients, metals, biological oxygen demand, and suspended solids based on US EPA standard methods (MWRD, 2010).

In accordance with the yearly single collection aspect of LTEF data, the three MWRD water samples per year were averaged to generate yearly means. Eleven water quality variables were obtained from MWRD database including fecal coliform (colony forming units per deciliter), dissolved oxygen (mg/l), five day biological oxygen demand (mg/l), total suspended solids (mg/l), ammonia (mg/l), total phosphorus (mg/l), pH, phenolic compounds (µg/l), total iron (mg/l), total manganese (mg/l), and total mercury (mg/l) (Table 2.1). Secchi disk visibility depth (cm) was obtained from LTEF data.

Individual LTEF sites, within the four reaches for which MWRD data was available, were associated with the nearest MWRD site. When associating the sites, preference was given to MWRD sites located upstream of LTEF sites. In two instances, one MWRD site was used for two separate, but proximal LTEF sites. Distances between LTEF and MWRD sites ranged from 0.08-4.67km with a mean of 1.96km. Following site association, the mean of MWRD variables were calculated for each reach to provide a single yearly value for each variable within each reach. The coefficient of variation and Pearson's correlation coefficient between each water quality variable and year were calculated in R 2.14.1 software. This allows for a determination of the relative magnitude of change in variables between reaches, and whether changes were directional through time.

Climate: National Climatic Data Center

We obtained regional climate data from National Oceanic and Atmospheric Administration's (NOAA) National Climatic Data Center (NCDC). Data was compiled from four sites with complete monthly data for the time period analyzed. Beginning upstream, sites were located in Chicago, Ottawa, Peru, and Peoria. Chicago climate data were used for Dresden reach, Ottawa for Marseilles reach, and Peru for Starved Rock reach. Due to the relatively large size of Peoria reach and the locations of two climate sites at the extreme ends of the reach, climate data from Peru and Peoria were averaged to better estimate climate throughout the entire reach. The approximate distance between

NCDC site locations and the center of the reach for which they were used varied; Dresden 50km, Marseilles 5km, Starved Rock 20km, and for Peoria both sites were approximately 40km from the center of the reach. The monthly data were converted to yearly climate variables for temperature and precipitation. This yielded sixteen climate variables, with eight corresponding to temperature and eight to precipitation (Table 2.1).

Choices of Modeling Approaches

A wide range of modeling approaches are available to examine assemblage-environment relationships (Austin et al., 2006; Dornelas et al., 2013; Gotelli et al., 2010). Data-mining methods are increasingly used for their capabilities of dealing with complex non-linear relationships and flexibility concerning assumptions (Allan et al., 2012; Hoang et al., 2013; Hochachka et al., 2007). One of these methods is Random Forests regression (RF) (Breiman, 2001). This method is an extension of Classification and Regression Trees (Breiman et al., 1984). Binary regression trees work through a process of recursive partitioning, which involves starting with all samples of the response variable and then recursively splitting the sample into two sub-groups (nodes) using a single predictor variable chosen from possible predictors to minimize average within-group variance. Methods generally used to determine the best split, therefore the best predictor variable, are based upon ordinary least squares regression or the sum of absolute deviations about a median (De'ath and Fabricius, 2000; Loh, 2011).

The advantages of regression trees include, no *a priori* knowledge is necessary in tree creation, ease of construction, ability to deal with missing data, able to handle a variety of response and predictor types (e.g. numeric, categorical), unaffected by monotonic transformations of explanatory variables, no assumptions of the distribution of predictor variables, and ease of interpretation (De'ath and Fabricius, 2000; Lewis, 2000). On the other hand, several disadvantages exist. Namely, they can miss additional predictor variables near the end of the tree resulting in terminal nodes with very few data points (“over-training”), miss synergistic interactions, and don’t fully capture the scale of a continuous factor (Cook and Goldman, 1984; Marshall, 2001). Finally, one of the largest disadvantages of regression trees is a high vulnerability to overfitting, which occurs when the tree too closely matches the training data (Khoshgoftaar et al., 2001). This leads to a tree which is accurate when applied to the training data, but much less so when new data is introduced. Even when not applied to prediction, overfitting can lead to inaccurate conclusions regarding the relationships between response and predictors, particularly with small sample sizes that may not be completely reflective of the ecosystem (Burnham and Anderson, 2002; O'Connor et al., 2004).

Random forest models are composed of multiple regression trees. Individual trees are built based upon discrete bootstrap samples using 2/3 of the response data. Each split within a tree is made using a defined number of predictors (M) randomly drawn from the pool of variables. For instance, if $M = 3$, at each node 3 variables are randomly selected from all available variables (p). The variable that best binarily splits the data is then selected for that node. Each tree is grown without restriction until terminal nodes can no longer be split, sometimes resulting in a single observation within terminal nodes. The tree is then applied to a portion (1/3) of response samples set aside for prediction (Breiman, 2001). By bootstrapping the response for each tree and randomly drawing from the predictor pool at each split within trees, one can build thousands of unique trees and then average the predictions across these trees.

The performance of a RF model can be evaluated with the percent of response variation explained (R^2). Significance of an individual predictor can be inferred with an importance value, which is estimated as the percent increase in model mean squared error (MSE) when the model is re-run with randomized values of that predictor. If a predictor is important, MSE will increase. If a predictor is not important, MSE may decrease or remain unchanged. Additionally, the effect of a predictor on the response variable can be investigated with partial dependence plots (Cutler et al., 2007). Partial dependence plots are generated by averaging out the effects of all but the focal predictor on the response. Fixed values across the range of the focal predictor variable are run through the RF model, and predicted response values are plotted against the fixed predictor values. These plots are inferentially limited to the evaluation of only the most important variables, and can be influenced by interactions (Friedman, 2001).

RF models overcome the disadvantage of overfitting through bagging, and averaging over many trees (Breiman, 2001). Over-training is overcome using multiple trees and small random subsets of predictors at each node. As a regression tree, RF permits the use of high dimensional data (the number of predictor variables is much greater than the number of response observations, p>>n), and interactions among variables (Genuer et al., 2010; Genuer et al., 2008). Though RF allows for interacting effects, it is difficult to assess interactions.

The one advantage, which single regression trees have over RF, is their ease of interpretation. Regression trees can be easily viewed as a whole, while the ensemble nature of RF makes examination of every tree unfeasible and meaningless. Although importance values and partial-dependence plots provide some insight into the relative importance of various variables to the prediction, overall RF does not offer sufficient information about the contribution of individual variables to the model. Multiple linear regression (MLR) was employed here to compliment RF for a better understanding of the relationship between fish assemblage structure, water quality, and climate variability.

MLR requires assumptions, which RF does not (normal distributions, linear relationships, homoscedasticity), and limits the number of predictor variables which can be used (Bingham et al., 2010). However, MLR has the advantage of letting the user predetermine models based upon ecological expectations. This allows pre-assigned combinations of predictors, and interacting effects. These model variations can then be compared against one another to aid in determination of which aspects of the environment an assemblage attribute is responding to most strongly. In our case use of MLR necessitated a modified logit transformation of the proportional response variables (S_p ,A_p, W_p, S_n, A_n, W_n) to bound predicted values between zero and one, $y_i = \ln(\frac{y_i+0.01}{1-y_i+0.01})$ (Warton and Hui, 2010). For more direct comparisons between RF and MLR, transformed response variables were used in both modeling approaches.

Model Calibrations

RF Regression

All seven logit-transformed response variables were modeled with RF using the 29-predictor variables (Table 2.1) for each reach separately, resulting in 28 total models. For each RF model, we used 5,000 trees, as many regression trees are necessary to stabilize estimates of variable importance (Liaw and Wiener, 2002). As stated, RF allows the user to select the number of randomly drawn variables available for selection at each node (*M*). We tried a range of *M* (1-p/2) for each model. Each level of *M* was run 10 times for each model. We selected the model that achieved the highest R^2 among all *M* values. If several models achieved similar R^2, the one with lowest *M* value was selected for further evaluation. Further investigation into models only occurred if the model was meaningful (explained > 10% of the variation in the response). Importance values of predictors for the model were recorded. The top five most important predictors, ones that most increased model MSE when randomized, were then investigated with partial-dependence plots. All RF modeling was carried out using package "randomForest" with R 2.14.1 software (Liaw and Wiener, 2002).

MLR

The seven logit transformed response variables were modeled for each reach individually. Our goal was to look for biological effects, so 15 competing models were constructed to represent biologically driven hypotheses using twelve predictor variables (Table 2.2). The pre-assignment of biological meaning aids in deciphering the causation of changes observed in the fish assemblages. MLR models were performed with R 2.14.1 software package "stats".

Model 1-2: Two models representing water clarity were used (*Secchi; Secchi+TSS*). Secchi depth provides a simple index of water clarity, and total suspended solids (TSS) is a primary factor affecting clarity (Borkman and Smayda, 1998; Preisendorfer, 1986). Clarity also directly affects predation and predator avoidance by prey (Brusven and Rose, 1981; Johnson and Hines, 1999; Shingles et al., 2005). Therefore, predator based attributes were expected to respond strongly to clarity models.

Model 3-9: Seven climate models were built based on temperature and precipitation. Three models were based solely on temperature, minimum temperature (*TMIN*), maximum temperature (*TMAX*), and both (*TMIN+TMAX*). As poikilothermic organisms, temperature is a principal factor affecting fish metabolic rates, which consequently affect growth rates (biomass), and fecundity (abundance) (Bullock, 1955). These three temperature models were expected to be related to both fish biomass and abundance. Four more models were based on precipitation, which not only affects flow, but also the inputs and levels of contaminants, nutrients, pH, temperature, terrestrial forage, and turbidity of rivers (Beven and Hornberger, 1982; Malmqvist and Rundle, 2002; Mason and MacDonald, 1982). These four models were based on monthly maximum precipitation (*PMAX*), which can capture the effects of heavy rainfall and flooding; minimum monthly precipitation (*PMIN*), which can capture light rainfall and drought; average rainfalls (*PMIN*), which can capture mean flow and interannual variation in total rainfall; and their overall combined effects (*PMAX+PMIN+PAVG*). These factors have individual and joint effects on fish diversity (Hughes and Gammon, 1987; Kawaguchi et al., 2003), so were expected to have possible corollary effects on all response variables.

Model 10: Fecal coliform has been found to be significantly correlated with numerous pollution variables associated with urban wastewater, including biological oxygen demand, turbidity, suspended solids, and nutrients (Sukias et al., 2001; Vis et al., 1998). The direct effect of sewage on the fish assemblages was modeled based on the interaction of fecal coliform and maximum precipitation (*FecCol×PMAX*), due to raw sewage discharged during heavy rainfall by the CSO systems employed by the cities of Chicago, Joliet, and Peoria. This model is expected to correlate with all assemblage response variables, as differences in the tolerance of fish species to several aspects of sewage pollution have a substantial effect on fish assemblage structure (Tsai, 1973).

Model 11: A single model representing three important nutrient variables (*NH3+P+Fe*) was conducted. Iron was included in the model due to its role as a dominant micronutrient, and its' significant impact on phosphorus bioavailability. Iron in the water column reacts with oxygen to form ferric oxide, which has a strong affinity for phosphate and renders both phosphorus and iron non-bioavailable (Buffle et al., 1989; Jones-Lee and Lee, 2005). Concentrations of these nutrients can be heavily influenced by urban and agricultural runoff, as well as natural geology. A literature review conducted by Wiejters et al. (2009) found associations between land use, nutrient concentrations, and fish diversity. Furthermore, the authors concluded that changes in fish assemblages were more sensitive to alterations than other organisms. In a long-term study conducted over a 17 year period, the fish assemblage responded to a marked decrease in nutrients, following improved sewage treatment, with increases in predatory fish abundance (Jeppesen et al., 1998). The nutrient model was expected to have a dominant effect on predatory biomass and abundance through trophic effects.

Model 12: Three main fish stressors were represented in a single model (*DO+NH3+TMAX*). Stress in fish results in increased oxygen demand, increased ventilation rates, and is associated with reduced growth (Bonga, 1997). Fish species vary greatly in their tolerance to individual stressors (Davis, 1975; Randall and Tsui, 2002).

Additionally, early life stages are more vulnerable to stressors than adults. For instance, low dissolved oxygen can cause mortality and deformities to eggs and larvae at concentrations tolerated by adults (Davis, 1975). Exposure to stressors also triggers longitudinal and vertical movements (Kramer, 1987). The combination of reduced growth, increased energy expenditures, decreased reproductive success, and movements because of stress has the potential to dramatically restructure fish assemblages in terms of diversity, biomass, and abundance.

Models 13-15: Finally, three global models were calibrated. The first was solely based on water quality variables (*DO+FecCol+NH3+P+Fe+Secchi+TSS*). The second was based on all climate variables used in previous models (*TMIN+TMAX+PAVG+PMAX+PMIN*). Lastly, a complete global model combining the effects of both water quality and climate was included (*TMIN+TMAX+PAVG+PMIN+ Domg+FecCol×PMAX+NH3+Phos+Fe+Secchi+TSS*). These models allow for evaluation of whether water quality, climate, or the combined effects of both had the most significant impact on fish assemblage structure.

Akaike Information Criterion adjusted for small sample size (AICc) was used to evaluate and rank MLR models, which has been shown to reduce overfitting when high-dimensional data are used (Bedrick and Tsai, 1994). AICc scores were calculated with package “AICcmodavg 1.27” using R 2.14.1 software (Mazerolle, 2013). Variable importance was calculated as the sum of model weights, based upon AICc scores, in which each variable was represented (Murray and Conner, 2009). To compensate for the unequal inclusion of variables into models, we divided the cumulative variable weight of a given variable by the number of models which included said variable (Kittle et al., 2008).

Results

Temporal Changes

There were differences observed in the response variables among reaches (Table 2.3). At least one of the two factors, sampling year and reach, was significant in all response variables (ANOVA, $p \leq 0.05$). Only the proportion of predator biomass (W_p) was found to be significantly different among years and reaches (Table 2.4). Some response variables displayed strong linear trends by year, while others varied less and were more stochastic (Table 2.3, Figure 2.2).

Species richness (S) increased in all reaches during the time period analyzed (Table 2.3). This increase was consistent for Peoria Reach, while other reaches were associated with varying degrees of fluctuation. The lowest value for S in all reaches, with the exception of Dresden Reach, occurred in 1990, the year following the severe 1988-1990 Midwest drought (Angel et al., 1992). In comparison, water quality variables during the drought remained relatively stable except total suspended solids (TSS) and, to a lesser extent, fecal coliform that increased considerably during the drought within the three reaches that had low S the following year (Figure 2.3).

All six proportional response variables (S_p , A_p, W_p, S_n, A_n, W_n) strongly increased over time in Dresden Reach ($0.69 \leq$ Pearson $r \leq 0.91$). These assemblage attributes also increased in Marseilles and Starved Rock Reach, but were associated with greater interannual variation. The changes in proportional abundance and biomass in these upper three reaches were largely driven by increases in native predatory centrarchid species, and decreases in non-native non-predatory common carp (*Cyprinus carpio*) (Figure 2.4). Largemouth bass (*Micropterus salmoides*) and bluegill (*Lepomis macrochirus*) drove the centrarchid biomass change in the Dresden and Marseilles Reaches. While these species also drove much of the biomass change in Starved Rock Reach, smallmouth bass (*Micropterus dolomieu*) also played a large role. Centrarchid abundance changes were driven predominantly by *L. macrochirus*, and green sunfish (*Lepomis cyanellus*). Much like biomass, *M. dolomieu* also contributed greatly to changes in abundance within Starved Rock Reach. The temporal change in proportional response variables for Peoria was disjointed, with some decreasing and others increasing through time (Table 2.3). Of these proportional response variables for Peoria Reach, only the proportion of native species (S_n) had a strong ($r = -0.62$) temporal trend. Unlike the clearer trends in the upper three reaches, there were no obvious temporal trends in the abundance or biomass of individual species in Peoria Reach.

Water quality variables had similar temporal trends among reaches (Table 2.5). Several variables noticeably increased across reaches, including dissolved oxygen, secchi visibility depth, phosphorus, pH, and iron. The concentrations of phenols, ammonia, and mercury decreased in all reaches. Fecal coliform decreased quite linearly over time in Dresden ($r = -0.88$), but much less so in remaining reaches ($-0.35 \leq r \leq -0.28$). Manganese and total suspended solids displayed little directional change throughout all reaches.

Modeling assemblage attributes-environment relationship with RF regression

When reporting the results of RF models we refer to models that explain greater than 10% of the observed variation in the response as “meaningful”. Ranking of model performance is based upon R^2, which RF reports as the percent variation explained in the 1/3 of samples retained for testing each tree. Of the 28 models (seven response variables × four reaches) attempted, 18 were meaningful ($11.4\% \leq R^2 \leq 73.3\%$; Table 2.6). Of these meaningful models, species richness (S) only performed well in Peoria Reach. All six proportional response variables performed best in the most upstream reach, Dresden ($42.3\% \leq R^2 \leq 73.3\%$), with the model for the proportion of native species abundance (A_n) performing best overall. There was a notable and steady decrease in model performance into downstream reaches for all proportional models. The best example of downstream decrease were models for W_p, the only response variable that accounted for some significant variation in all four reaches (R^2: Dresden 68.0%, Marseilles 21.4%, Starved Rock 18.9%, Peoria 15.6%).

The five most important variables for each of the eighteen models were identified through ranking the variables, which caused the largest increase in model mean squared error (MSE) when only the said variable was randomized and the model rerun. The most important predictors varied among responses (Table 2.7). Water quality and climate variables accounted for 80% and 20% of the top five predictors, respectively, across models. Climate variable were mainly selected for the downstream reach, Peoria. Most of the top five variables had a clear positive or negative, approximately linear relationship with all fish assemblage variables (Table 2.8, Figure 2.5). Dissolved oxygen, secchi depth, iron, and ammonia occurred in ≥ 50% of all meaningful models (Table 2.8). The most important climate variables were the number of days the temperature was at or below freezing (DX32) and the average mean monthly rainfall (PAVG) with three total occurrences each (16.7% of models) (Table 2.8).

For the RF model of species richness (S) in Peoria Reach, the only meaningful model for this response variable, the most important predictor variables were secchi, iron, dissolved oxygen, yearly mean of extreme minimum monthly temperatures (TMIN), and phenols (Table 2.7).

Seven of 12 models for native fish (3 response variables × 4 reaches) were meaningful. The proportion of native species (S_n) model was only meaningful for Dresden Reach (Table 2.6) with secchi, dissolved oxygen, fecal coliform, yearly mean of extreme minimum monthly temperatures (TMIN), and phenols ranked as the most important (Table 2.7). Three of four models for proportion of native species abundance (A_n) were meaningful with secchi, dissolved oxygen, and ammonia ranked within the top 5 variables of two models, and fecal coliform (27.76% increase in MSE) of high importance to Dresden Reach (Table 2.7). The models for proportion of native species biomass (W_n) were also meaningful for three reaches, with pH, ammonia, dissolved oxygen, and phosphorus occurring in two of the three models (Table 2.7). Secchi had the highest importance value among W_n models (40.80%), but was only included into the top five variables in one reach, Dresden.

Ten of 12 predator-based models (3 response variables × 4 reaches) were meaningful (Table 2.6). Iron was included into the top five variables of all 3 meaningful proportion of predator species (S_p) models, and was followed by secchi and dissolved oxygen, which were included in the top five variables of 2 of the 3 (Table 2.7).

Dissolved oxygen and phosphorus were prominent in all three proportion of predator abundance (A_p) models, and secchi was included in two of three (Table 2.7). Ammonia and iron were included into the top five variables of one model for A_p, but had relatively high importance in those models. Ammonia increased MSE by 30.24% for Dresden, and iron increased MSE by 26.11% for Marseilles. The proportion of predator biomass (W_p) models explained differing amounts of variation (15.6%-68.0%) for all four reaches. Ammonia was included in all W_p models, and secchi, iron, dissolved oxygen, and pH were each included in two of four (Table 2.7).

We also examined importance from a reach perspective (Table 2.9). Six models were meaningful for Dresden, and three variables were included into all six; secchi depth (35.62% - 43.94%), dissolved oxygen (14.73% - 28.4%), and fecal coliform (12.87% - 27.76%). The most prominent variables for Marseilles Reach were iron (26.11% - 43.76%), secchi depth (9.75% - 44.97), and pH (8.95% - 37.05%). Three variables were included into both Starved Rock models; dissolved oxygen (17.62% - 23.09%), iron (12.45% - 12.7%), and mercury (6.36% - 11.69%). Three of the five meaningful models for Peoria included iron (8.87% - 22.12%), pH (8.59% - 14.47%), and PAVG (6.2% - 11.95%) within the top five variables.

MLR

The best models resulting from MLR were selected using AICc scores. If multiple models competed with a Δ AICc of ≤ 2, or a Δ AICc between 2-3 with a higher k than the Δ AICc = 0 model, the model with the highest R^2 was selected. Of the 28 best models, 12 were water clarity models containing either secchi alone or the combined effects of secchi and TSS (Table 2.10). Stress and three of four precipitation models accounted for the majority of the remaining models, five each (Table 2.10).

Similar to the RF analyses, best models tend to show a gradual decline in model performance downstream with the exception of species richness. All proportional assemblage models performed best, assessed by R^2, for Dresden Reach (Table 2.10). A graph of additive variable importance for the models of biomass and abundance reveals that the two proportional abundance response variables (A_p, A_n) are responding most strongly to water clarity in the form of secchi depth, and the two biomass responses (W_p, W_n) are responding most strongly to variables which affect primary and secondary productivity (Figure 2.6); including high temperatures (TMAX), dissolved oxygen (DO), and nutrient levels (NH3, Fe, and P).

Discussion

Our results indicated that proportional changes in the IRW's fish assemblages since 1983 are strongly associated with improvements in water quality throughout the waterway. Three variables in particular appear to have had the largest influence, dissolved oxygen, ammonia, and water clarity. Across RF and MLR models, there is also a consistent decrease in model performance from upstream to downstream with climate variables playing an increasing role in the most downstream reach of this study (Peoria). These results imply that Chicago area effluents were the primary factor for fish assemblage changes in the upper reaches, but impacts declined downstream. Presumably, this is caused by dilution, sequestration, and settling effects as the river moves downstream increasing in volume and encountering dams. In comparison, Koel and

Sparks (2002) examined pre-alteration hydrological data (1879-1899) against post-alteration hydrological and fish data (1957-1998), and concluded hydrological alterations and interannual variability in the timing and magnitude of precipitation significantly altered the Illinois River fish assemblages. Our study does not necessarily conflict with theirs, because different time frames were involved. Hydrologic alterations (e.g. channelization) were persistent between 1983-2010, so are unlikely to be the primary cause of observed fish assemblage changes in our study. Using precipitation data as a proxy for flow variability, we've found that water quality likely played a greater role in the structuring of the fish assemblages in the four IRW reaches from 1983-2010 than interannual flow variability.

Our analyses provided insights into the relative importance of individual water quality variables. Improvements in sewage treatment leading to improved water quality was noted as early as 1975 (Koel and Sparks, 2002). CSOs acting as the primary structuring influence on the Upper IRW fish assemblages are suggested by the nature of the predictor variables identified as most important. CSOs are known to affect dissolved oxygen, water clarity, and nutrient levels in large rivers (Even et al., 2004; Field et al., 1997; Iannuzzi et al., 1997). We found fecal coliform was included within the top five RF variables among all models for Dresden. This result should not be surprising, because Dresden Reach is nearest Chicago, and the prevalence of fecal coliform has been shown to be a good indicator of sewage overflow intensity (Vis et al., 1998). West et al. (2011) found fecal coliform correlated with phosphorous and reduced water clarity. However, we found relatively high levels of phosphorus have been maintained, though clarity increased and fecal coliform declined. Our study indicates that while improved sewage treatment and overflow reduction have had significant impacts on clarity, ammonia, and dissolved oxygen, they have had little impact on phosphorus concentrations. David and Gentry (2000) estimated 70% of the phosphorus entering the IRW between 1980-1997 was attributable to sewage. However, the authors also concluded there is likely a vast store of phosphorus in soils, and possibly river sediment, due to unused fractions caused by over fertilization during the 1960s and 70s. They surmise that the combination of phosphorus stores and continued sewage inputs will contribute a large fraction of phosphorus to the river system for the near future.

The counterintuitive inverse relationship between clarity and phosphorus observed also warrants further discussion. We have a few possible explanations for the increase in water clarity amid high concentrations of phosphorus. First, phosphorus may not be acting as a limiting nutrient to primary production in the IRW. It is possible the nitrogen reductions observed limited primary production, and alleviated water clarity degradation. Second, concentrations of cations, especially iron (III), are rendering orthophosphate non-bioavailable. Moreover, this relationship may have been enhanced by the increased oxygen levels (Correll, 1998; Jones-Lee and Lee, 2005). Third, the increased clarity was principally driven by factors unrelated to primary production (e.g. reduced sediment loads). Fourth, our use of total phosphorus does not adequately reflect the fraction that is bioavailable as orthophosphate, so is not a good indicator of IRW trophic state. Lastly, and most likely, it is combination of the four possible explanations. More studies are necessary to determine specific sources of phosphorus in the IRW, and what impact these concentrations may have on the ecosystem.

The effects of water quality and climate variables varied among assemblage attributes and reaches. Species richness models consistently performed the poorest among models, with the exception of Peoria Reach. This is possibly due to the higher number of sampling sites in Peoria Reach, which increases the detection probability of rare species, and allows for a more accurate depiction of the increase in species richness. It appeared predator and native proportions in Peoria were much more stable (Figure 2.2). This could indicate Peoria Reach is more functionally stable, and explain why proportional responses did not model as well (Figure 2.2). However, though it may be more functionally stable, the increase in species richness does indicate that environmental conditions were generally improving. The poor performances of species richness models in other reaches may be attributable to a possibly poor estimation of species richness due to the relatively small number of sampling sites in those reaches. However, other assemblage attributes strongly responded to the predictors in those reaches despite the smaller number of sampling sites (Tables 2.6, 2.10). This result highlights the importance of examining multiple aspects of assemblage structure.

One of the most interesting findings was that biomass and abundance responded to different predictor variables in MLR. Biomass was mostly responsive to nutrients, oxygen, and high temperatures, an observation that is biologically meaningful as these factors influence growth potential (Jobling, 1981; Polis et al., 1997). In comparison, abundance mainly responded to clarity measures in our analyses. Fish sampling occurs following the warmest most productive months of the year, summer, thus the effect of higher temperatures and increased primary productivity can be detected in fish biomass without a time lag. It should be stated that in the beginning stages of our study, both one and two-year time lags in climate variables were applied to modeling approaches, resulting in decreased model performance as evaluated by R^2. Random Forest regression did not reveal these differences in abundance and biomass environment relationships, which shows the utility of applying complementary analyses to assess biologically significance Abundances' response to clarity could be due to increased electrofishing sampling efficiency in clearer water (Sanders, 1992). However, we believe it is due to clarities ability to connote overall water quality as an indicator of sediment loads, trophic state, and habitat variables (e.g. flow, light penetrance) (Davies-Colley and Smith, 2001; Megard et al., 1980).

Overall, the results of this study indicate observed increases of predatory and native species are due to water quality (i.e. sewage treatment) improvements. In 1972 when the CWA became law, municipalities were required to meet effluent requirements primarily aimed at controlling biodegradable matter, and bacterial pollution (Copeland, 1999). Chicago initiated the Deep Tunnel Project shortly after passage of the CWA. Phase 1 of the project has been partially operational since 1980, was completed in 2006, and stores 2.3 billion gallons for secondary wastewater treatment (Scalise and Fitzpatrick, 2012). Treatment of raw sewage, which would have previously been released directly into the IRW, is likely the source of observed water quality improvements. Phase 2 will increase storage volume to 17.5 billion gallons, and some reservoirs have already been completed. Decreases in raw sewage releases should continue to have a significant effect on river biota until, and after, completion of all reservoirs in 2029. It is possible to marginally predict the changes of predatory and native fishes using our model with the expected future improvements in water quality. However, we suspect the positive effects of further water quality improvement will be constrained by existing hydrological and geomorphological alteration. Continued monitoring of fish assemblages and water quality is essential to test this hypothesis and serve the need for restoring the IRW ecosystem.

Consistent with our study, fish diversity in other rivers has been found to increase following pollution mitigation efforts. The Rivers Trent and Thames in England, and the Rhine in mainland Europe displayed increases in species richness and changes to assemblage structure in terms of native and predator fishes following water quality mitigation efforts (Araújo et al., 2000; Fedorenkova et al., 2013; Langford et al., 2012; Plum and Schulte-Wülwer-Leidig, 2012; Raat, 2001). Results from these rivers agree with our findings that increased dissolved oxygen and decreased ammonia are primary drivers of fish recovery in water quality degraded systems.

The lower occurrence and importance of climate variables, along with relatively high R^2 values for some models suggest that water quality is a main driver of fish assemblage structure in the IRW. Though inclusion of climate data, precipitation and temperature, did benefit our analyses. Climate variable importance was generally lower than that of water quality, but permutation of climate variables in RF generally had a negative effect on model performance. This indicates climate data did explain a portion of observed variation. Inclusion of climate data likely reduced stochastic variability in the fish assemblage attributes associated with yearly precipitation and temperature variations.

Results of our study have several important implications for future researchers and managers. 1) Single category (guild) proportional functional responses can capture fish assemblage response to environmental alterations. 2) Combining biological knowledge and multiple modeling approaches aids in interpretation. 3) Water quality improvements alone can greatly enhance fish assemblages. This is especially pertinent in large rivers, in which it is economically infeasible to return to “natural” hydrological state due to shipping and agricultural needs. 4) Through cooperation of multiple agencies, gaps in data needed for ecological assessment can be filled.

In conclusion, our findings support the hypothesis that water quality had a major influence on the structure of IRW fish assemblages. The value of complementary environmental data should not be underestimated, and researches should make every effort to seek data from outside sources for inclusion in analyses. This is especially pertinent to long-term studies where the original intent was monitoring alone. Without the ability to assign possible causation to biological surveys, ecology reverts to a purely descriptive science. Current researchers should evaluate the efficacy of collecting their own environmental data when conducting long-term surveys for later analyses. Further, researchers with existing biological survey data should actively seek complimentary environmental data for analyses.

References

Allan, JD, Yuan, LL, Black, P, Stockton, T, Davies, PE, Magierowski, RH and Read, SM. 2012. Investigating the relationships between environmental stressors and stream condition using Bayesian belief networks. *Freshwater Biology* **57**: 58-73.

Angel, JR, Bender, GM, Durgunolu, A, Hollinger, SE, Jones, DMA, Knapp, HV, Kunkel, KE, McConkey, SA, Portis, DH, Ramamurthy, GS, Scott, RW, Semonin, RG, Singh, KP, Wehrmann, HA and Wendland, WM. 1992. The 1988-1989 drought in Illinois: causes, dimensions, and impacts. Illinois State Water Survey, Champaign, Illinois.

Araújo, FG, Williams, WP and Bailey, RG. 2000. Fish assemblages as indicators of water quality in the middle Thames estuary, England (1980–1989). *Estuaries* **23**: 305-317.

Austin, M, Belbin, L, Meyers, J, Doherty, M and Luoto, M. 2006. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. *Ecological Modelling* **199**: 197-216.

Bedrick, EJ and Tsai, C-L. 1994. Model selection for multivariate regression in small samples. *Biometrics* **50**: 226-231.

Beven, KJ and Hornberger, GM. 1982. Assessing the effect of spatial pattern of precipitation in modeling stream flow hydrographs. *Journal of the American Water Resources Association* **18**: 823-829.

Bingham, NH, Bingham, N and Fry, JM. 2010. *Regression: linear models in statistics*. Springer, London, UK.

Bonga, SW. 1997. The stress response in fish. *Physiological reviews* **77**: 591-625.

Borkman, DG and Smayda, TJ. 1998. Long-term trends in water clarity revealed by secchi-disk measurements in lower Narragansett Bay. *ICES Journal of Marine Science* **55**: 668-679.

Boulton, AJ. 1999. An overview of river health assessment: philosophies, practice, problems and prognosis. *Freshwater Biology* **41**: 469-479.

Breiman, L. 2001. Random Forests. *Machine Learning* **45**: 5-32.

Breiman, L, Friedman, JH, Olshen, RA and Stone, CJ. 1984. *Classification and regression trees*. Chapman & Hall, Belmont, CA.

Brusven, M and Rose, S. 1981. Influence of substrate composition and suspended sediment on insect predation by the torrent sculpin, *Cottus rhotheus*. *Canadian Journal of Fisheries and Aquatic Sciences* **38**: 1444-1448.

Buffle, J, De Vitre, R, Perret, D and Leppard, GG. 1989. Physico-chemical characteristics of a colloidal iron phosphate species formed at the oxic-anoxic interface of a eutrophic lake. *Geochimica et Cosmochimica Acta* **53**: 399-408.

Bullock, TH. 1955. Compensation for temperature in the metabolism and activity of poikilotherms. *Biological Reviews* **30**: 311-342.

Bunn, SE and Arthington, AH. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management* **30**: 492-507.

Burnham, KP and Anderson, DR. 2002. *Model selection and multi-model inference: a practical information-theoretic approach*. Springer, New York, NY.

Butts, TA, Evans, RL and Lin, S. 1975. Water quality features of the Upper Illinois Waterway. Illinois State Water Survey, Champaign, IL.

Byers, JE. 2002. Impact of non-indigenous species on natives enhanced by anthropogenic alteration of selection regimes. *Oikos* **97**: 449-458.

Cook, EF and Goldman, L. 1984. Empiric comparison of multivariate analytic techniques: advantages and disadvantages of recursive partitioning analysis. *Journal of chronic diseases* **37**: 721-731.

Copeland, C. 1999. *Clean Water Act: a summary of the law*. Congressional Research Service, Library of Congress, Washington, DC.

Correll, DL. 1998. The role of phosphorus in the eutrophication of receiving waters: a review. *Journal of environmental quality* **27**: 261-266.

Crawley, MJ. 2012. *The R book*. John Wiley & Sons.

Cutler, DR, Edwards Jr, TC, Beard, KH, Cutler, A, Hess, KT, Gibson, J and Lawler, JJ. 2007. Random forests for classification in ecology. *Ecology* **88**: 2783-2792.

David, MB and Gentry, LE. 2000. Anthropogenic inputs of nitrogen and phosphorus and riverine export for Illinois, USA. *Journal of environmental quality* **29**: 494-508.

Davies-Colley, RJ and Smith, DG. 2001. Turbidity, suspended sediment, and water clarity: a review. *Journal of the American Water Resources Association* **37**: 1085-1101.

Davis, JC. 1975. Minimal dissolved oxygen requirements of aquatic life with emphasis on Canadian species: a review. *Journal of the Fisheries Board of Canada* **32**: 2295-2332.

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

Donohoe, M. 2003. Causes and health consequences of environmental degradation and social injustice. *Social Science & Medicine* **56**: 573-587.

Dornelas, M, Magurran, AE, Buckland, ST, Chao, A, Chazdon, RL, Colwell, RK, Curtis, T, Gaston, KJ, Gotelli, NJ and Kosnik, MA. 2013. Quantifying temporal change in biodiversity: challenges and opportunities. *Proceedings of the Royal Society: Biological Sciences* **280**: rspb20121931

Douglas, D and McCreanor, J. 1990. The Big River, Co Louth, Ireland: a case study in recovery. *Annales de Limnologie.* **26**: 73-79.

Douglas, I. 1990. Sediment transfer and siltation. In: The Earth as transformed by human action. Press Syndicate of the University of Cambridge, New York, NY, pp 215-234

Dudgeon, D, Arthington, AH, Gessner, MO, Kawabata, Z-I, Knowler, DJ, Lévêque, C, Naiman, RJ, Prieur-Richard, A-H, Soto, D, Stiassny, MLJ and Sullivan, CA. 2005. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* **81**: 163.

Elliott, JM. 1990. The need for long-term investigations in ecology and the contribution of the Freshwater Biological Association. *Freshwater Biology* **23**: 1-5.

Even, S, Poulin, M, Mouchel, J-M, Seidl, M and Servais, P. 2004. Modelling oxygen deficits in the Seine River downstream of combined sewer overflows. *Ecological Modelling* **173**: 177-196.

Fedorenkova, A, Vonk, JA, Breure, AM, Hendriks, AJ and Leuven, RS. 2013. Tolerance of native and non-native fish species to chemical stress: a case study for the River Rhine. *Aquatic Invasions* **8**: 231-241.

Field, R, Pitt, RE, Fan, C-Y, Heaney, JP, Stinson, MK, DeGuida, RN, Perdek, JM, Borst, M and Hsu, KF. 1997. Urban wet-weather flows. *Water Environment Research* **69**: 426-444.

Friedman, JH. 2001. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*: 1189-1232.

Frimpong, EA and Angermeier, PL. 2009. Fish Traits: a database of ecological and life-history traits of freshwater fishes of the United States. *Fisheries* **34**: 487-495.

Galat, DL and Zweimüller, I. 2001. Conserving large-river fishes: is the highway analogy an appropriate paradigm? *Journal of the North American Benthological Society* **20**: 266-279.

Genuer, R, Poggi, J-M and Tuleau-Malot, C. 2010. Variable selection using random forests. *Pattern Recognition Letters* **31**: 2225-2236.

Genuer, R, Poggi, J-M and Tuleau, C. 2008. Random Forests: some methodological insights. *Research report INRIA Saclay* **RR-6729**: 1-32.

Gotelli, NJ, Dorazio, RM, Ellison, AM and Grossman, GD. 2010. Detecting temporal trends in species assemblages with bootstrapping procedures and hierarchical models. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**: 3621-3631.

Guégan, J-F, Lek, S and Oberdorff, T. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* **391**: 382-384.

Harrell, HL. 1978. Response of the Devil's River (Texas) fish community to flooding. *Copeia* **1978**: 60-68.

Harris, J and Silveira, R. 1999. Large-scale assessments of river health using an Index of Biotic Integrity with low-diversity fish communities. *Freshwater Biology* **41**: 235-252.

Hay-Jahans, C. 2011. *An R companion to linear statistical models*. CRC Press, Boca Raton, FL.

Hoang, TH, Mouton, A, Lock, K, De Pauw, N and Goethals, PL. 2013. Integrating data-driven ecological models in an expert-based decision support system for water management in the Du River Basin (Vietnam). *Environmental Monitoring and Assessment* **185**: 631-642.

Hochachka, WM, Caruana, R, Fink, D, Munson, A, Riedewald, M, Sorokina, D and Kelling, S. 2007. Data-mining discovery of pattern and process in ecological systems. *The Journal of Wildlife Management* **71**: 2427-2437.

Hughes, RM and Gammon, JR. 1987. Longitudinal changes in fish assemblages and water quality in the Willamette River, Oregon. *Transactions of the American Fisheries Society* **116**: 196-209.

Iannuzzi, T, Huntley, S, Schmidt, C, Finley, B, McNutt, R and Burton, S. 1997. Combined sewer overflows (CSOs) as sources of sediment contamination in the lower Passaic River, New Jersey. I. Priority pollutants and inorganic chemicals. *Chemosphere* **34**: 213-231.

Jeppesen, E, Søndergaard, M, Jensen, JP, Mortensen, E, Hansen, A-M and Jørgensen, T. 1998. Cascading trophic interactions from fish to bacteria and nutrients after reduced sewage loading: an 18-year study of a shallow hypertrophic lake. *Ecosystems* **1**: 250-267.

Jobling, M. 1981. The influences of feeding on the metabolic rate of fishes: a short review. *Journal of Fish Biology* **18**: 385-400.

Johnson, JE and Hines, RT. 1999. Effect of suspended sediment on vulnerability of young razorback suckers to predation. *Transactions of the American Fisheries Society* **128**: 648-655.

Jones-Lee, A and Lee, GF. 2005. Role of iron chemistry in controlling the release of pollutants from resuspended sediments. *Remediation Journal* **16**: 33-41.

Kawaguchi, Y, Taniguchi, Y and Nakano, S. 2003. Terrestrial invertebrate inputs determine the local abundance of stream fishes in a forested stream. *Ecology* **84**: 701-708.

Khoshgoftaar, TM, Allen, EB and Deng, J. 2001. Controlling overfitting in software quality models: experiments with regression trees and classification. *Seventh International Software Metrics Symposium (METRICS 2001).* London, UK.

Kittle, AM, Fryxell, JM, Desy, GE and Hamr, J. 2008. The scale-dependent impact of wolf predation risk on resource selection by three sympatric ungulates. *Oecologia* **157**: 163-175.

Koel, TM and Sparks, RE. 2002. Historical patterns of river stage and fish communities as criteria for operations of dams on the Illinois River. *River Research and Applications* **18**: 3-19.

Kramer, DL. 1987. Dissolved oxygen and fish behavior. *Environmental Biology of Fishes* **18**: 81-92.

Langford, T, Worthington, T, Shaw, P, Kemp, P, Woolgar, C, Ferguson, A, Harding, P and Ottewell, D. 2012. The unnatural history of the River Trent: 50 years of ecological recovery. In: *River Conservation and Management*. John Wiley & Sons, Ltd, New Delhi, India, pp 261-272

Langford, TE, Shaw, PJ, Howard, SR, Ferguson, AJ, Ottewell, D and Eley, R. 2010. Ecological recovery in a river polluted to its sources: the River Tame in the English Midlands. In: *Ecology of Industrial Pollution*. Cambridge University Press, Cambridge, UK, pp 255-275

Lerczak, TV, Sparks, RE and Blodgett, DK. 1994. The long-term Illinois River fish population monitoring program project F-101-R. Illinois Natural History Survey, Champaign, IL.

Levêque, C, Oberdorff, T, Paugy, D, Stiassny, M and Tedesco, P. 2008. Global diversity of fish (Pisces) in freshwater. In: *Freshwater Animal Diversity Assessment*. Springer, pp 545-567

Lewis, RJ. 2000. An introduction to classification and regression tree (CART) analysis. *Annual Meeting of the Society for Academic Emergency Medicine.* San Francisco, California

Lian, Y, You, J, Sparks, R and Demissie, M. 2012. Impact of human activities to hydrologic alterations on the Illinois River. *Journal of Hydrologic Engineering* **17**: 537-546.

Liaw, A and Wiener, M. 2002. Classification and regression by randomForest. *R News* **2/3**: 18-22.

Loh, W-Y. 2011. Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **1**: 14-23.

Lubinski, KS, Sparks, RE and Jahn, LA. 1974. The development of toxicity indices for assessing the quality of the Illinois River. Illinois Natural History Survey, Champaign, IL.

Malmqvist, Br and Rundle, S. 2002. Threats to the running water ecosystems of the world. *Environmental conservation* **29**: 134-153.

Marshall, RJ. 2001. The use of classification and regression trees in clinical epidemiology. *Journal of Clinical Epidemiology* **54**: 603-609.

Mason, C and MacDonald, S. 1982. The input of terrestrial invertebrates from tree canopies to a stream. *Freshwater Biology* **12**: 305-311.

Mazerolle, MJ. 2013. AICcmodavg: model selection and multimodel inference based on AICc version 1.27 (R package)

McClelland, MA, Pegg, MA and Spier, TW. 2006. Longitudinal patterns of the Illinois Waterway fish community. *Journal of Freshwater Ecology* **21**: 91-99.

McClelland, MA and Sass, GG. 2012. Assessing fish collections from random and fixed site sampling methods on the Illinois River. *Journal of Freshwater Ecology* **27**: 325-333.

McClelland, MA, Sass, GG, Cook, TR, Irons, KS, Michaels, NN, O'Hara, TM and Smith, CS. 2012. The long-term Illinois River fish population monitoring program. *Fisheries* **37**: 340-350.

Megard, RO, Settles, JC, Boyer, HA and Combs, WS, Jr. 1980. Light, secchi disks, and trophic states. *Limnology and Oceanography* **25**: 373-377.

Meybeck, M. 2003. Global analysis of river systems: from Earth system controls to Anthropocene syndromes. *Philosophical Transactions of the Royal Society of London: Biological Sciences* **358**: 1935-1955.

Minns, CK, Kelso, JR and Randall, RG. 1996. Detecting the response of fish to habitat alterations in freshwater ecosystems. *Canadian Journal of Fisheries and Aquatic Sciences* **53**: 403-414.

Moore, M, Romano, SP and Cook, T. 2010. Synthesis of upper Mississippi River system submersed and emergent aquatic vegetation: past, present, and future. *Hydrobiologia* **640**: 103-114.

Murray, K and Conner, MM. 2009. Methods to quantify variable importance: implications for the analysis of noisy ecological data. *Ecology* **90**: 348-355.

MWRD. 2010. Illinois Waterway monitoring project quality assurance project plan. Monitoring and Research Department, Chicago, IL.

O'Connor, RJ, Wagner, TL and Sodhi, N. 2004. A test of a regression-tree model of species distribution. *The Auk* **121**: 604-609.

Oberdorff, T, Tedesco, PA, Hugueny, B, Leprieur, F, Beauchard, O, Brosse, S and Dürr, HH. 2011. Global and regional patterns in riverine fish species richness: a review. *International Journal of Ecology* **2011**: Article ID 967631

Pegg, MA and McClelland, MA. 2004. Spatial and temporal patterns in fish communities along the Illinois River. *Ecology of Freshwater Fish* **13**: 125-135.

Plum, N and Schulte-Wülwer-Leidig, A. 2012. From a sewer into a living river: the Rhine between Sandoz and Salmon. *Hydrobiologia*: 1-12.

Poff, NL, Olden, JD, Merritt, DM and Pepin, DM. 2007. Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences* **104**: 5732-5737.

Polis, GA, Anderson, WB and Holt, RD. 1997. Toward an integration of landscape and food web ecology: the dynamics of spatially subsidized food webs. *Annual Review of Ecology and Systematics*: 289-316.

Preisendorfer, RW. 1986. Secchi disk science: visual optics of natural waters. *Limnology and Oceanography*: 909-926.

Raat, AJ. 2001. Ecological rehabilitation of the Dutch part of the River Rhine with special attention to the fish. *Regulated Rivers: Research & Management* **17**: 131-144.

Raibley, PT, Irons, KS, O'Hara, TM, Blodgett, KD and Sparks, RE. 1997. Winter habitats used by largemouth bass in the Illinois River, a large river-floodplain ecosystem. *North American Journal of Fisheries Management* **17**: 401-412.

Randall, D and Tsui, T. 2002. Ammonia toxicity in fish. *Marine Pollution Bulletin* **45**: 17-23.

Romano, P, Ranzani, R and Tecchiati, F. 1992. Water reclamation of the Po River in the Turin area. *Water Science & Technology* **26**: 2579-2582.

Sanders, RE. 1992. Day versus night electrofishing catches from near-shore waters of the Ohio and Muskingum Rivers. *The Ohio Journal of Science* **92.3**: 51-59.

Scalise, C and Fitzpatrick, K. 2012. Chicago deep tunnel design and construction. *Structures Congress 2012.* Chicago, IL.

Shingles, A, McKenzie, D, Claireaux, G and Domenici, P. 2005. Reflex cardioventilatory responses to hypoxia in the flathead gray mullet (*Mugil cephalus*) and their behavioral modulation by perceived threat of predation and water turbidity. *Physiological and Biochemical Zoology* **78**: 744-755.

Sparks, RE. 2010. Forty years of science and management on the Upper Mississippi River: an analysis of the past and a view of the future. *Hydrobiologia* **640**: 3-15.

Sparks, RE and Lerczak, TV. 1993. Recent trends in the Illinois River indicated by fish populations. Illinois Natural History Survey, Champaign, IL.

Sparks, RE and Sandusky, MJ. 1981. Identification of factors responsible for decreased production of fish food organisms in the Illinois and Mississippi Rivers. Illinois Natural History Survey, Champaign, IL.

Strayer, D, Glitzenstein, JS, Jones, CG, Kolasa, J, Likens, GE, McDonnell, MJ, Parker, GG, Pickett, ST and Arboretum, MFC. 1986. Long-term ecological studies: an illustrated account of their design, operation, and importance to ecology. Institute of Ecosystem Studies, Millbrook, NY.

Sukias, JP, Tanner, CC, Davies-Colley, RJ, Nagels, JW and Wolters, R. 2001. Algal abundance, organic matter, and physico-chemical characteristics of dairy farm facultative ponds: implications for treatment performance. *New Zealand Journal of Agricultural Research* **44**: 279-296.

Theriot, C and Tzoumis, K. 2005. Deep tunnels and fried fish: tracing the legacy of human interventions on the Chicago River. *Golden Gate University of Law Review* **35**: 377.

Tsai, C-F. 1973. Water quality and fish life below sewage outfalls. *Transactions of the American Fisheries Society* **102**: 281-292.

Turner, MG, Collins, SL, Lugo, AL, Magnuson, JJ, Rupp, TS and Swanson, FJ. 2003. Disturbance dynamics and ecological response: the contribution of long-term ecological research. *BioScience* **53**: 46-56.

Vis, C, Hudon, C, Cattaneo, A and Pinel-Alloul, B. 1998. Periphyton as an indicator of water quality in the St Lawrence River (Quebec, Canada). *Environmental Pollution* **101**: 13-24.

Warton, DI and Hui, FKC. 2010. The arcsine is asinine: the analysis of proportions in ecology. *Ecology* **92**: 3-10.

Weijters, M, Janse, J, Alkemade, R and Verhoeven, J. 2009. Quantifying the effect of catchment land use and water nutrient concentrations on freshwater river and stream biodiversity. *Aquatic Conservation: Marine and Freshwater Ecosystems* **19**: 104-112.

West, BM, Liggit, P, Clemans, DL and Francoeur, SN. 2011. Antibiotic resistance, gene transfer, and water quality patterns observed in waterways near CAFO farms and wastewater treatment facilities. *Water, Air, & Soil Pollution* **217**: 473-489.

Wootton, JT and Power, ME. 1993. Productivity, consumers, and the structure of a river food chain. *Proceedings of the National Academy of Sciences* **90**: 1384-1387.

Figures and Tables: Chapter 2

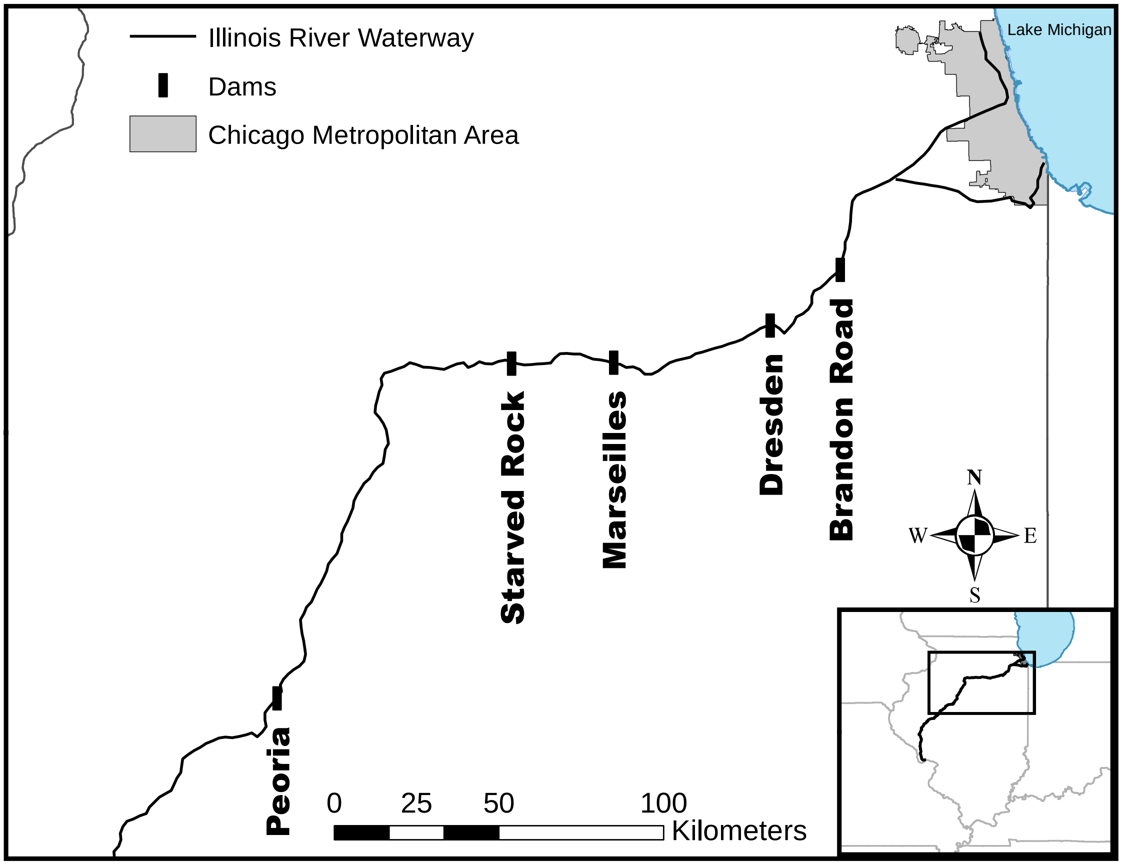


Figure 2.1: Map of our study area. The Upper Illinois River Waterway is divided into four reaches defined by lock and dam complexes. Reaches are named for pools formed by the dams located at their downstream extent.

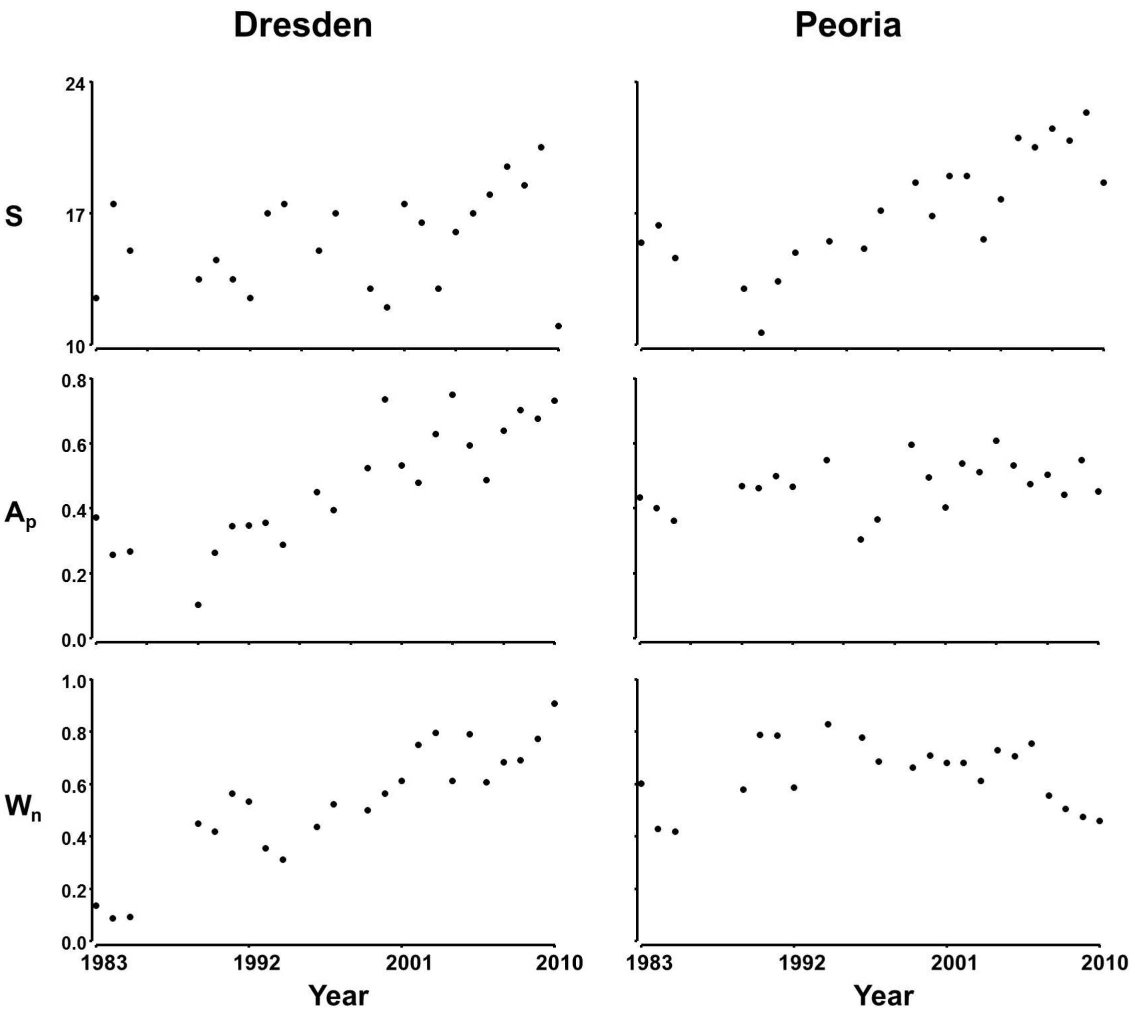


Figure 2.2: Long-term changes of three selected response variables in Dresden and Peoria reaches based on annual sampling (S=species richness, A_p=proportion predator abundance, W_n=proportion native biomass).

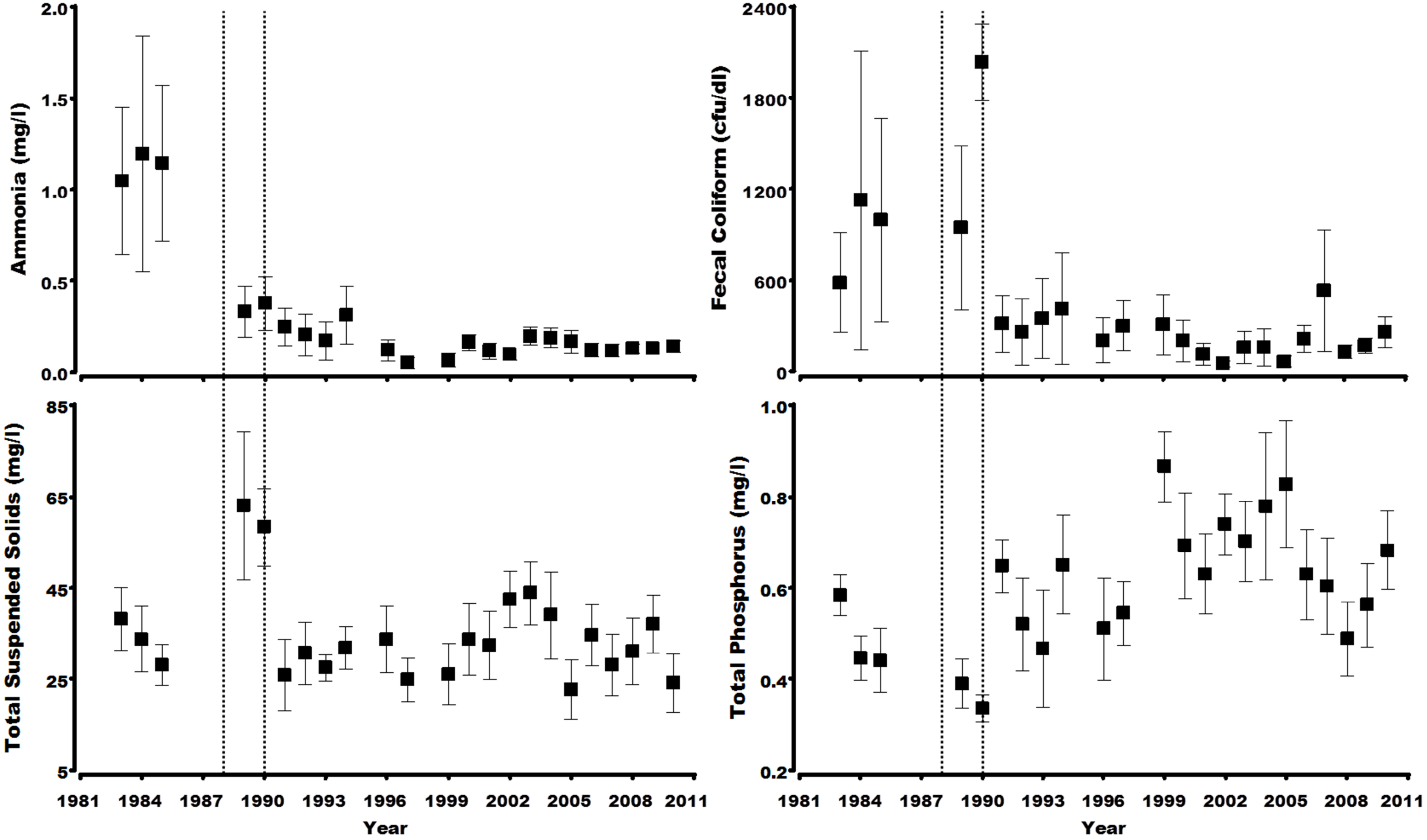


Figure 2.3: Temporal trends in ammonia, fecal coliform, suspended solids, and phosphorus across all four reaches with standard error. Dashed line demarcates the severe 1988-1990 Midwest drought years.

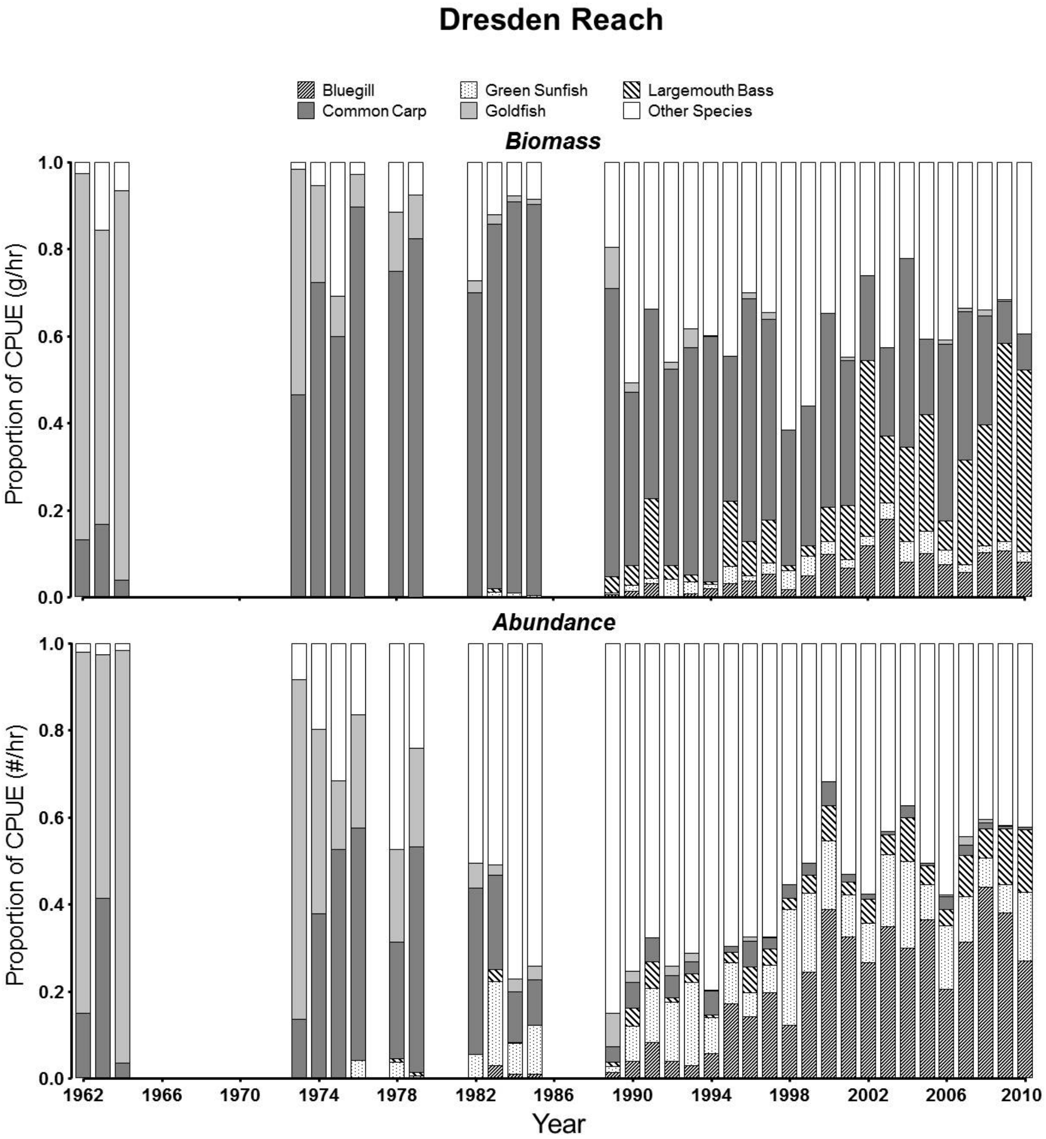


Figure 2.4: Proportional changes for selected species in Dresden Reach during 1957-2010.

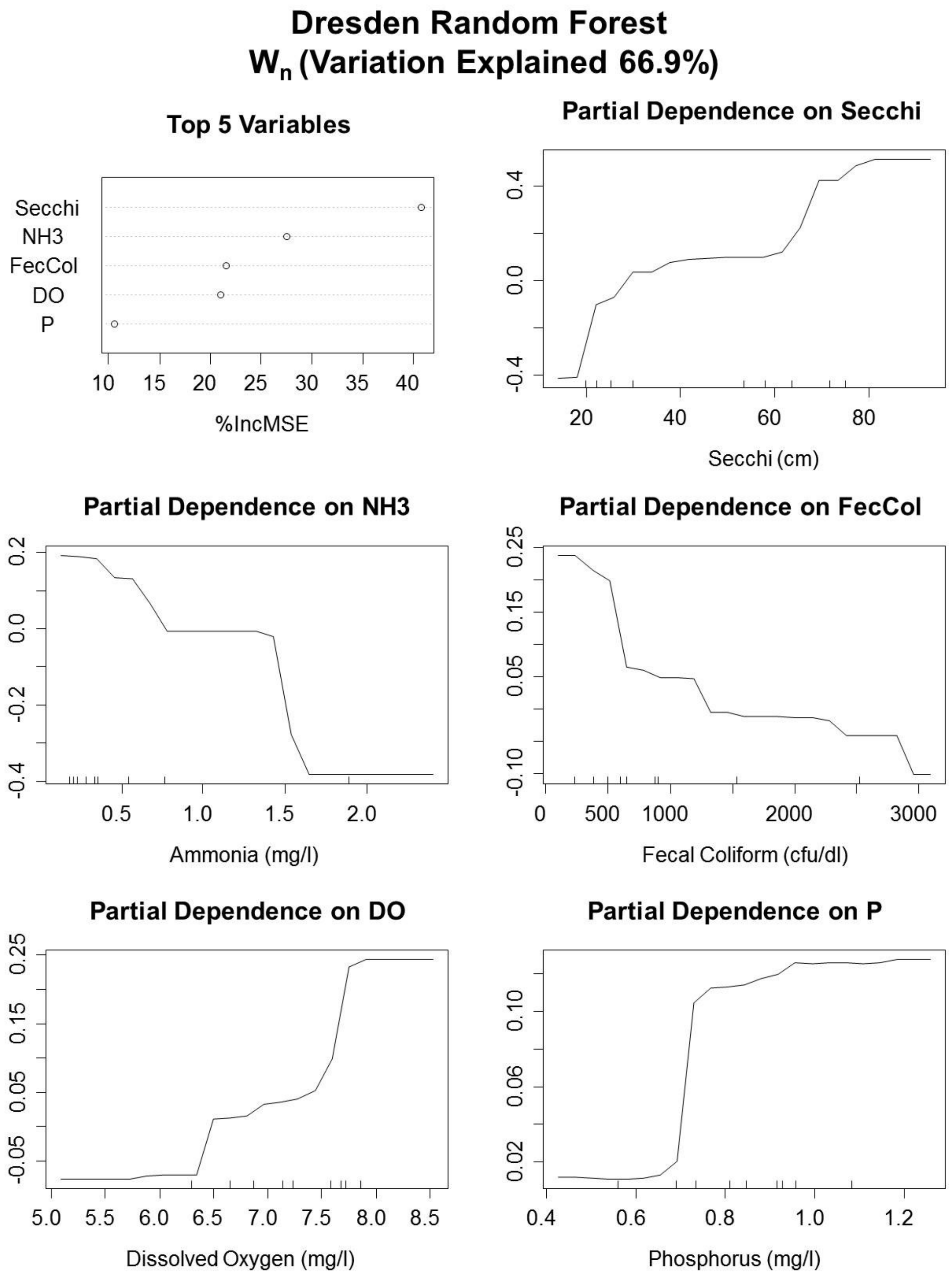


Figure 2.5: Importance values of the top five variables in random forests regression of the proportion native biomass (W_n) in Dresden Reach, and corresponding partial dependence plots. The y-axis of partial dependence plots represents the response variable, but the scale is not informative.

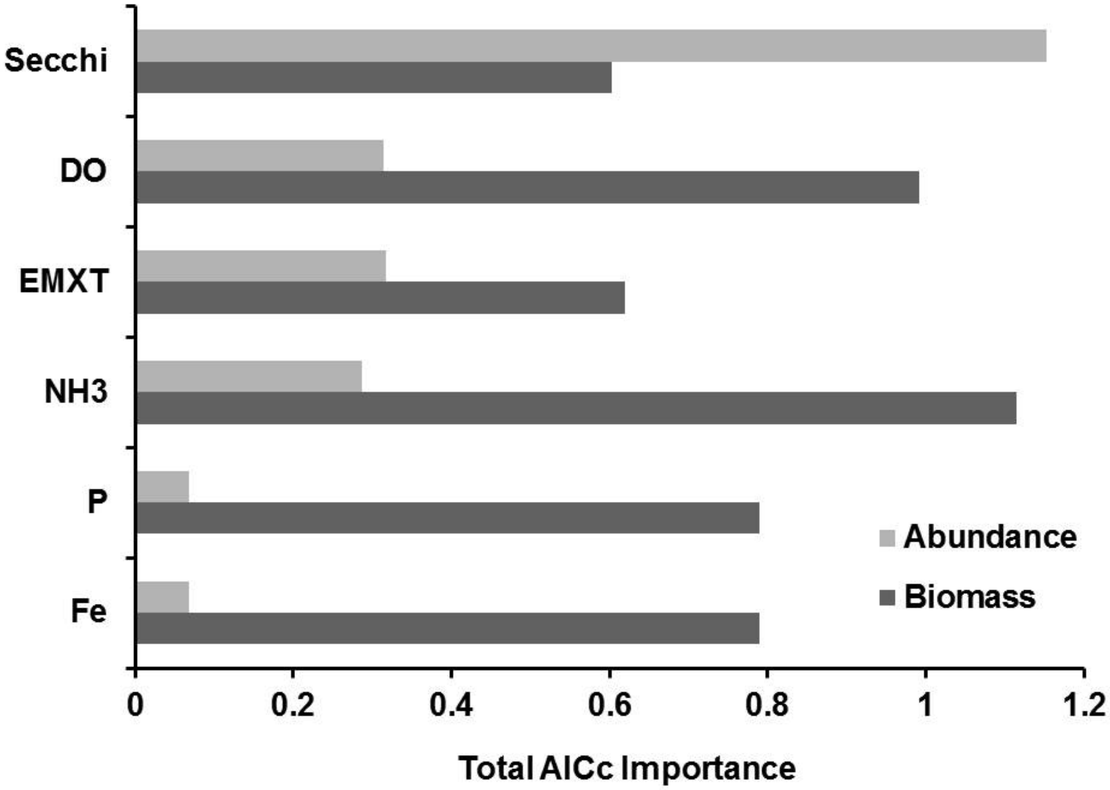


Figure 2.6: Summed AICc variable weight of multiple linear regression modeling adjusted for number of possible models in which variables are included across reaches. Biomass component includes proportion predator biomass and proportion native biomass response variable models. Abundance includes proportion predator abundance and proportion native abundance response variable models.

Table 2.1: Environmental predictor variables used in random forest regression models.

Variable	Definition	Median	Range	Units
Water Quality				
Secchi	Secchi Visibility Depth	30.7	10.6 - 93	cm
FecCol	Fecal Coliform	118.6	10.8 - 3093.3	cfu/dl
DO	Dissolved Oxygen	7.96	5.09 - 11.18	mg/l
BOD5	5-Day Biological Oxygen Demand	3.31	0 - 9.5	mg/l
TSS	Total Suspended Solids	28.7	11.2 - 102.8	mg/l
NH3	Ammonia	0.14	0 - 2.4	mg/l
P	Phosphorus	0.55	0.29 - 1.26	mg/l
pH	pH	7.7	6.7 - 8.7	
Phen	Phenolic Compounds	5.36	0.94 - 30.40	µg/l
Fe	Iron	0.6	0.2 - 8.9	mg/l
Mn	Manganese	0.05	0.03 - 0.41	mg/l
Hg	Mercury	0.07	0 - 0.25	mg/l
Climate				
DP01	Days with Precipitation ≥ 0.1"	70	50 – 93	Days
DP05	Days with Precipitation ≥ 0.5"	24	12 – 37	Days
DP10	Days with Precipitation ≥ 1.0"	8	0 – 15	Days
DT00	Days with Min Temp < 0 F	9	1 – 28	Days
DT90	Days with Max Temp ≥ 90 F	17	1 – 50	Days
DX32	Days with Max Temp < 32 F	45	15 – 71	Days
PMAX	Yearly Mean Monthly Max Rainfall	17.3	10.3 - 36.2	cm
PTOT	Total Yearly Precipitation	94.0	56.1 - 137.3	cm
PAVG	Yearly Mean Monthly Mean Rainfall	7.8	4.7 - 11.4	cm
PMIN	Yearly Mean Monthly Min Rainfall	1.8	0 - 4.2	cm
TSNW	Total Yearly Snowfall	67	14 – 173	cm
TMIN	Yearly Mean of Extreme Min Monthly Temps	-3.8	-7 - -0.2	°C
TMAX	Yearly Mean of Extreme Max Monthly Temps	25.4	23.2 - 27.5	°C
TMNA	Yearly Mean of Mean Monthly Min Temps	-10.9	-15.7 - -4	°C
TMXA	Yearly Mean of Mean Monthly Max Temps	15.8	13.9 - 18.3	°C
TAVG	Yearly Mean of Mean Monthly Temps	10.4	8.0 - 12.2	°C

Table 2.2: Multiple linear regression hypotheses and accompanying variables.

Hypothesis	Variables
Secchi alone	Secchi
Temperature	TMIN + TMAX
Max Temperature	TMAX
Min Temperature	TMIN
Max Precipitation	PMAX
Min Precipitation	PMIN
Avg Precipitation	PAVG
Clarity	Secchi + TSS
Combined Sewage Overflow	FecCol × PMAX
Nutrients	NH3 + Phos + Fe
Precipitation	PAVG + PMAX + PMIN
Stress	DO + NH3 + TMAX
Water Quality	DO + FecCol + NH3 + Phos + Fe + Secchi + TSS
Climate	TMIN + TMAX + PAVG + PMAX + PMIN
Global (all variables)	TMIN + TMAX + PAVG + PMIN + Domg + FecCol × PMAX + NH3 + Phos + Fe + Secchi + TSS

Table 2.3: Summary of response variables used in models. The year 1984 is missing from Starved Rock, and 1993 is missing from Peoria. Standard deviation is labeled as S.D., and the coefficient of variation is labeled as CV. Pearson correlation coefficient for each response with year is provided as r. (* indicates strong temporal trends >0.50)

Reach	Response	n	Mean	Range	S.D.	CV	r
Dresden	**S**	23	15.57	11.00-20.50	2.6	16.7	0.33
	S_p	23	0.53	0.40-0.70	0.08	14.13	0.79*
	A_p	23	0.48	0.10-0.75	0.18	38.78	0.85*
	W_p	23	0.43	0.04-0.83	0.23	54.56	0.91*
	S_n	23	0.88	0.79-0.96	0.05	5.62	0.69*
	A_n	23	0.93	0.76-0.99	0.07	7.32	0.87*
	W_n	23	0.53	0.09-0.91	0.22	42.42	0.90*
Marseilles	**S**	23	14.59	11.67-21.33	2.62	17.97	0.53*
	S_p	23	0.51	0.36-0.60	0.07	13.85	0.43
	A_p	23	0.36	0.12-0.59	0.13	37.46	0.61*
	W_p	23	0.32	0.06-0.63	0.15	47.59	0.73*
	S_n	23	0.92	0.88-0.98	0.02	2.67	0.15
	A_n	23	0.95	0.83-1.00	0.04	4.59	0.69*
	W_n	23	0.55	0.12-0.90	0.21	38.38	0.76*
Starved Rock	**S**	22	14.66	8.00-23.50	4.13	28.17	0.64*
	S_p	22	0.51	0.33-0.64	0.09	17.73	0.56*
	A_p	22	0.24	0.05-0.47	0.14	55.37	0.30
	W_p	22	0.32	0.01-0.78	0.18	56.92	0.63*
	S_n	22	0.94	0.89-1.00	0.03	3.06	0.20
	A_n	22	0.98	0.87-1.00	0.03	3.19	0.53*
	W_n	22	0.67	0.24-1.00	0.23	34.8	0.50
Peoria	**S**	22	17.17	10.63-22.38	3.05	17.79	0.79*
	S_p	22	0.56	0.51-0.64	0.03	5.9	-0.17
	A_p	22	0.47	0.30-0.61	0.08	16.11	0.40
	W_p	22	0.31	0.18-0.39	0.05	15.85	0.20
	S_n	22	0.9	0.84-0.96	0.03	3.32	-0.62*
	A_n	22	0.89	0.78-0.97	0.06	6.78	-0.02
	W_n	22	0.64	0.42-0.83	0.12	19.43	0.01

Table 2.4: Results of ANOVA for differences in response variables among reaches and years.

Response	**Reach P**	**Year P**
S	0.0525	<0.0001
$\mathbf{S_p}$	0.1683	0.0001
$\mathbf{A_p}$	<0.0001	0.076
$\mathbf{W_p}$	0.0155	0.0007
$\mathbf{S_n}$	<0.0001	0.6721
$\mathbf{A_n}$	<0.0001	0.1285
$\mathbf{W_n}$	0.1072	0.0003

Table 2.5: Table showing the coefficient of variation, and Pearson correlation coefficient with year (r) for each water quality variable by Reach. (* indicates a strong temporal trend > 0.50 or < -0.50)

Variable	**Reach**							
	Dresden		**Marseilles**		**Starved Rock**		**Peoria**	
	CV	**r**	**CV**	**r**	**CV**	**r**	**CV**	**r**
Secchi	51.79	0.88*	55.43	0.81*	48.90	0.75*	43.12	0.80*
FecCol	88.05	-0.69*	176.28	-0.35	235.15	-0.38	211.20	-0.28
DO	10.31	0.68*	12.70	0.65*	12.91	0.65*	12.38	0.53*
BOD5	60.97	-0.29	59.76	-0.11	65.01	-0.06	43.03	0.23
TSS	22.44	-0.31	45.61	-0.19	55.68	-0.29	17.71	-0.13
NH3	113.37	-0.79*	136.98	-0.70*	128.96	-0.63*	55.39	-0.51*
P	24.82	0.62*	23.82	0.46	26.30	0.42	25.31	0.35
pH	3.16	0.56*	5.34	0.73*	5.38	0.63*	6.20	0.72*
Phen	83.14	-0.78*	93.23	-0.73*	99.67	-0.69*	88.02	-0.76*
Fe	77.04	0.18	137.05	0.23	155.16	0.14	127.08	0.29
Mn	67.13	-0.06	100.72	0.05	117.02	0.05	73.27	0.07
Hg	77.95	-0.37	83.48	-0.45	83.01	-0.47	83.70	-0.42

Table 2.6: Percent variation explained by Random Forest models. Dash indicates model was unable to explain any of the observed variance.

Reach	**S**	**S_p**	**A_p**	**W_p**	**S_n**	**A_n**	**W_n**
Dresden	-	50.5%	61.2%	68.0%	42.3%	73.3%	66.9%
Marseilles	-	32.0%	14.1%	21.4%	-	37.8%	42.8%
Starved Rock	-	11.4%	-	18.9%	-	-	-
Peoria	42.4%	-	13.8%	15.6%	-	15.1%	21.2%

Table 2.7: Summary of RF variable importance by response.

Response	Models	Variable	Occurrences	Mean	Range
Species Richness (S)	1	Secchi	1	37.92%	
		Fe	1	22.12%	
		DO	1	14.74%	
		TMIN	1	11.68%	
		Phen	1	11.25%	
Proportion of Predator Species (S_p)	3	Fe	3	23.53%	12.70% - 43.76%
		Secchi	2	28.39%	13.15% - 43.63%
		DO	2	16.17%	14.73% - 17.62%
		tss	1	21.45%	
		FecCol	1	13.50%	
		P	1	12.79%	
		TSNW	1	12.31%	
		DX32	1	11.81%	
		Hg	1	11.69%	
		TMXA	1	8.88%	
		Phen	1	2.98%	
Proportion of Predator Species Abundance (A_p)	3	DO	3	15.43%	7.88% - 19.38%
		P	3	9.70%	5.49% - 12.64%
		Secchi	2	25.90%	13.56% - 38.24%
		NH3	1	30.24%	
		Fe	1	26.11%	
		FecCol	1	12.87%	
		DX32	1	10.74%	
		TMXA	1	9.72%	
		pH	1	8.59%	
		DT00	1	7.59%	
Proportion of Predator Species Biomass (W_p)	4	NH3	4	14.68%	3.28% - 26.22%
		Secchi	2	24.48%	9.75% - 39.22%
		Fe	2	20.72%	12.45% - 28.98%
		DO	2	19.59%	16.08% - 23.09%
		pH	2	13.92%	13.44% - 14.41%
		BOD	1	24.93%	
		TMAX	1	23.12%	
		P	1	21.78%	
		FecCol	1	20.23%	
		Phen	1	17.55%	
		Hg	1	6.36%	
		PAVG	1	6.20%	
		DP10	1	6.19%	

Table 2.7: Continued

Response	Models	Variable	Occurrences	Mean	Range
Proportion of Native Species (S_n)	1	Secchi	1	43.94%	
		DO	1	23.04%	
		FecCol	1	21.31%	
		TMIN	1	6.81%	
		Phen	1	6.50%	
Proportion of Native Species Abundance (A_n)	3	Secchi	2	40.29%	35.61% - 44.97%
		DO	2	20.72%	13.03% - 28.39%
		NH3	2	17.51%	10.47% - 24.55%
		FecCol	1	27.76%	
		BOD	1	14.67%	
		DP10	1	12.90%	
		PTOT	1	11.68%	
		DX32	1	11.14%	
		PAVG	1	10.70%	
		Hg	1	9.34%	
		pH	1	8.95%	
		Fe	1	8.87%	
Proportion of Native Species Biomass (W_n)	3	pH	2	25.76%	14.47% - 37.05%
		NH3	2	22.93%	18.29% - 27.57%
		DO	2	15.33%	9.58% - 21.08%
		P	2	14.31%	10.56% - 18.06%
		Secchi	1	40.80%	
		FecCol	1	21.61%	
		TMAX	1	13.14%	
		Phen	1	12.90%	
		PTOT	1	12.48%	
		PAVG	1	11.95%	
		Fe	1	11.59%	

Table 2.8: Summary of variable importance and performance for the 20 individual predictor variables across the 18 Random Forest models, in terms of occurrences in the top five important variables. Direction of relationship between the responses and predictors inferred from partial dependence plots.

Variable	**Total**	**% Occurrence**	**% Positive**	**% Negative**	**% Unimodal**
DO	13	72.2	100	0	0
Secchi	11	61.1	100	0	0
Fe	9	50	77.8	22.2	0
NH3	9	50	0	100	0
P	7	38.9	85.7	0	14.3
FecCol	6	33.3	0	100	0
pH	6	33.3	100	0	0
Phen	5	27.8	0	100	0
Hg	3	16.7	0	100	0
DX32	3	16.7	66.7	33.3	0
PAVG	3	16.7	0	100	0
BOD	2	11.1	50	0	50
TMIN	2	11.1	100	0	0
TMAX	2	11.1	100	0	0
TMXA	2	11.1	50	50	0
DP10	2	11.1	0	100	0
PTOT	2	11.1	0	100	0
DT00	1	5.6	0	100	0
TSNW	1	5.6	100	0	0
TSS	1	5.6	100	0	0

Table 2.9: Summary of RF variable importance by reach.

Reach	Models	Variable	Occurrences	Mean	Range
Dresden	**6**	Secchi	6	40.24%	35.62% - 43.94%
		DO	6	20.45%	14.73% - 28.40%
		FecCol	6	19.55%	12.87% - 27.76%
		NH3	4	27.15%	24.55% - 30.24%
		P	3	14.99%	10.56% - 21.78%
		BOD	1	14.67%	
		Fe	1	14.12%	
		TSNW	1	12.31%	
		TMIN	1	6.81%	
		Phenols	1	6.50%	
Marseilles	**5**	Fe	3	32.95%	26.11% - 43.76%
		Secchi	3	22.76%	9.75% - 44.97%
		pH	3	20.13%	8.95% - 37.05%
		DO	3	13.89%	9.58% - 19.05%
		NH3	3	12.85%	9.77% - 18.29%
		Phenols	3	11.15%	2.98% - 17.55%
		P	2	11.78%	5.49% - 18.06%
		DX32	2	11.28%	10.74% - 11.81%
		TSS	1	21.45%	
		Hg	1	9.34%	
		TMXA	1	8.88%	
Starved Rock	**2**	DO	2	20.35%	17.62% - 23.09%
		Fe	2	12.58%	12.45% - 12.70%
		Hg	2	9.02%	6.36% - 11.69%
		BOD	1	24.93%	
		Secchi	1	13.15%	
		P	1	12.79%	
		NH3	1	3.28%	
Peoria	**5**	Fe	3	14.19%	8.87% - 22.12%
		pH	3	12.17%	8.59% - 14.47%
		PAVG	3	9.62%	6.20% - 11.95%
		TMAX	2	18.13%	13.14% - 23.12%
		PTOT	2	12.08%	11.68% - 12.48%
		DO	2	11.31%	7.88% - 14.74%
		DP10	2	9.55%	6.19% - 12.90%
		Secchi	1	37.92%	
		NH3	1	19.46%	
		TMIN	1	11.68%	
		Phenols	1	11.25%	
		DX32	1	11.14%	
		P	1	10.95%	
		TMXA	1	9.72%	
		DT00	1	7.59%	

Table 2.10: Summary of the best MLR models for each response at each reach. K = number of regressor variables, ∆ AICc = the difference between the chosen model and the model with the lowest AICc score, R^2 = coefficient of determination for model, $adjR^2$ = coefficient of determination adjusted for number of regressors.

Response	Reach	Best Model	K	∆ AICc	$R^2/adjR^2$
Species Richness (S)	Dresden	Min Precip	3	0	0.14/0.10
	Marseilles	Secchi	3	0	0.28/0.25
	Starved Rock	Min Precip	3	0	0.15/0.11
	Peoria	Clarity	4	0	0.75/0.73
Proportion of Predator Species (S_p)	Dresden	Clarity	4	2.87	0.68/0.65
	Marseilles	Clarity	4	0	0.41/0.35
	Starved Rock	Clarity	4	2.44	0.37/0.31
	Peoria	Precipitation	5	0.47	0.32/0.21
Proportion of Predator Species Abundance (A_p)	Dresden	Clarity	4	0.23	0.63/0.59
	Marseilles	Stress	5	0	0.52/0.44
	Starved Rock	Stress	5	1	0.31/0.19
	Peoria	Nutrients	5	0.03	0.35/0.24
Proportion of Predator Species Biomass (W_p)	Dresden	WQ	9	0	0.91/0.88
	Marseilles	Nutrients	5	0	0.45/0.36
	Starved Rock	Stress	5	0	0.36/0.26
	Peoria	Stress	5	0	0.64/0.58
Proportion of Native Species (S_n)	Dresden	Clarity	4	2.64	0.66/0.63
	Marseilles	Min Precip	3	0	0.18/0.14
	Starved Rock	Avg Precip	3	0	0.19/0.15
	Peoria	Clarity	4	0.88	0.39/0.33
Proportion of Native Species Abundance (A_n)	Dresden	Secchi	3	0	0.78/0.79
	Marseilles	Clarity	4	1.22	0.45/0.39
	Starved Rock	Secchi	3	0	0.39/0.36
	Peoria	MaxTemp	3	0	0.20/0.16
Proportion of Native Species Biomass (W_n)	Dresden	WQ	9	0	0.91/0.87
	Marseilles	Nutrients	5	0	0.60/0.54
	Starved Rock	Secchi	3	0	0.18/0.14
	Peoria	Stress	5	2.86	0.48/0.39

Chapter 3

The Impact of Water Quality and Climate on the Habitat and Foraging Functional Diversity of Fish Assemblages

Abstract

The impacts of human disturbances on taxonomic diversity (i.e. species richness and evenness) have been examined in numerous studies. Functional diversity is emerging as an important dimension of biodiversity. It is defined as the diversity of biological traits that directly affect ecosystem functions (e.g. productivity, decomposition), and can be measured with a variety of indices. Functional diversity assessments have become increasingly common in terrestrial plant literature. However, few functional diversity studies have focused on fish assemblages, particularly in terms of functional responses to water quality changes. We assessed the long-term changes of fish assemblage functional diversity in the upper 207 km of the Illinois River Waterway (IRW) between 1983 and 2010 in relation to water quality and interannual climate variability. This study incorporates two functional diversity indices, Functional Evenness (FEve) and Functional Dispersion (FDis), which quantify separate functional structure components. The former describes the regularity with which the multidimensional trait space is filled, whereas the latter measures the spread of species away from the weighted center of trait space. We used two sets of traits, foraging and habitat preference to estimate these separate aspects of functional diversity. Abundance or biomass of individual fish species were used to weight the two indices (FEve, FDis). Random forests (RF) regression was used to assess the effects of water quality and climate variables on these indices. RF models explained up to 64.7% of the observed variation in

index values. We found water quality variables to be the more important factors for functional diversity, and biomass-based indices responded to water quality changes more strongly than abundance-based indices. These findings indicate that water quality and, to a lesser extent, climate strongly affected the foraging and habitat-based functional structure of the fish assemblages in the IRW.

Introduction

Biodiversity can be examined at multiple hierarchical levels, from genes to ecosystems (Bellard et al., 2012; Loreau, 2010). Much of biodiversity research and conservation efforts have focused on taxonomic diversity at the species level. Recently studies have begun to shift the focus toward functional diversity (FD). FD is a measurement of the variety and variability of organismal traits that influence one or more aspects of the functioning of an ecosystem (Tilman, 2001). Traits used for estimating FD include morphological, physiological, and behavioral characteristics that determine how an organism interacts with different features of its environment (Ernst et al., 2006; Flynn et al., 2009). This interaction, in turn, helps define an organism's functional role in processes such as nutrient cycling, productivity, resilience, and invasion resistance (Cadotte et al., 2011; Naeem et al., 2000; Tilman, 2001).

Use of FD indices are fairly new, but have potential to produce long sought after general ecological principles regarding how communities are structured by their environment (McGill et al., 2006). Thus far, the majority of FD studies have focused on terrestrial plant communities, but at least five studies have examined FD of fish assemblages (Mason et al., 2008; Pease et al., 2012; Pool et al., 2010; Schleuter et al., 2012; Villéger et al., 2010). These five studies have assessed overall FD using a combination of traits related to habitat and foraging, and found climate to be a primary driver of spatial FD patterns in fish assemblages within river basins and across regions. Villéger (2010) and Pease (2012) included some water quality data into their analyses, but no detailed toxin or nutrients data. Furthermore, these studies solely focused on species abundances for density measurements. However, body size differs greatly among species and individuals, thus abundance and biomass weighting have the potential to provide drastically different results (Cohen et al., 2003; Dauer et al., 1993). Evidence also suggests a stronger relationship between biomass and ecosystem function (Cardinale et al., 2007).

There is a fundamental need to increase our understanding of the effects human activities have on fish assemblage FD. Anthropogenic disturbances have caused global declines in the taxonomic diversity of river fishes (Bunn and Arthington, 2002; Dudgeon et al., 2006; Meybeck, 2003; Meybeck, 2004). It is estimated that extinction rates for North American freshwater fauna may be as much as 4% per decade, which is 5 times higher than extinction rates of the terrestrial fauna and as rapid as rates from tropical forests (Ricciardi and Rasmussen, 1999; Sala et al., 2000). However, how human disturbances, particularly water quality degradation, affect functional diversity has not been evaluated extensively. One such ecosystem in which fishes have been impacted by anthropogenic disturbance is the Illinois River Waterway (IRW).

Water Quality in the IRW is heavily influenced by municipal and industrial effluents, especially from the Chicago metropolitan area (Sparks and Sandusky, 1981). Degraded water quality led to extreme loss of fish species, with upper river segments affected most strongly (Sparks and Lerczak, 1993). Enhanced sewage treatment, industrial effluent regulations, and mitigation of untreated combined sewage overflow releases have greatly improved water quality since the 1970s (Sparks and Lerczak, 1993; Theriot and Tzoumis, 2005). During this time, an increasing trend has been observed in species richness and the abundance of predatory game species coincident with improvements in water quality (Koel and Sparks, 2002; Lerczak et al., 1994; Pegg and McClelland, 2004).

Fish assemblages of the IRW have been monitored nearly continuously since 1957, and have yielded a wealth of information regarding recent changes to the fish assemblages. However, previous studies of the IRW fish assemblages have been limited by a lack of sufficient water quality data, and functional analyses have been limited to a guild level approach (Chapter 2; McClelland et al., 2006; Pegg and McClelland, 2004). A guild level approach is useful, but much information is lost when analyses are confined to a single trait. This study seeks to fill the knowledge gap through use of FD indices that take into account multiple traits, and use of complementary water quality data containing multiple variables.

Habitat and foraging traits can be used to calculate FD jointly or separately. We chose the second approach to compare these two different aspects of FD (habitat, foraging). Physical habitat is known to be important to riverine fish diversity (Guégan et al., 1998). However, physical habitat is difficult to assess habitats in large rivers, and meant these data were not available for our analyses. Further, we assumed habitat had less effect on fish diversity than did water quality, because all major hydrological modifications were in place before 1940 (Lian et al., 2012) and fish were sampled at fixed sites. Some FD indices also have the ability to incorporate abundance or biomass as weighting. We wanted to compare the effectiveness of weighting indices with biomass over abundance, and determine if the density unit used as a weight altered fish assemblage FD responses to the environment.

The goal of this study is to examine the importance of water quality and climate in the functional structure of IRW fish assemblages. We examined the upper 207km of the IRW for the period 1983-2010 for which water quality and climate data were available (Figure 3.1). To accomplish our goal, we 1) compared the responses of foraging and habitat based FD to water quality changes and interannual climate variability, and 2) determine if biomass or abundance based FD changes correspond superiorly to our suite of environmental variables. Based upon the assumption that physical habitat has changed minimally at sampling locations, we hypothesize foraging-based FD will show stronger correlations with water quality. We further expect models using biomass-weighted indices will perform superiorly to abundance-weighted indices, as there is evidence of stronger relationships between fish biomass and FD (Carey and Wahl, 2011).

Methods

Data Sources

Fish Assemblage Data

Fish assemblage data were obtained from the Long-Term Fish Population Monitoring Program of the Illinois River, otherwise known as the Long-term Electrofishing project (LTEF). Standardized one-hour AC electrofishing was conducted at 26 fixed locations in the IRW. Sites are located longitudinally from the Lower Des Plaines River to the Lower Illinois River near its confluence with the Mississippi River. These sites are sampled once annually within a six week period from late August to early October. Water level and temperature sampling restrictions are followed to reduce the probability of fishes being dispersed into the floodplain and migrating to overwintering habitat (Pegg and McClelland, 2004).

Fish data was standardized into catch per unit effort (CPUE) for each species, as number per hour (CPUEn) and grams per hour (CPUEw). The sum of each species CPUEn and CPUEw was calculated by year for the four upper most reaches separately (Dresden, Marseilles, Starved Rock, Peoria). Reaches are defined as pools formed by five lock and dam complexes. Starting upriver, the Dresden Reach begins at Brandon Road lock and dam, and ends Dresden lock and dam. Next, Marseilles Reach stretches downstream to Marseilles lock and dam. The Starved Rock Reach continues downstream to Starved Rock lock and dam. Peoria Reach ends at Peoria lock and dam (Figure 3.1). Numbers of sampling sites varied among reaches, reflecting the distance between lock and dam complexes; Dresden contained two sites, Marseilles three sites, Starved Rock two sites, and Peoria eight sites.

Reach level division of the river has been well established by several previous studies (Koel and Sparks, 2002; McClelland et al., 2006; McClelland and Sass, 2012; McClelland et al., 2012; Pegg and McClelland, 2004; Sparks, 2010; Sparks and Lerczak, 1993). These studies demonstrated significant differences between the hydrological regime, elevation, slope, adjoining land cover, geology, chemistry, and fish diversity of the reaches (Butts et al., 1975; Koel and Sparks, 2002; Lian et al., 2012; McClelland et al., 2006; Pegg and McClelland, 2004). The dams create a physical barrier to fish passage, nutrient spiraling, and create clear divisions. Moreover, fish survey sites were occasionally sampled unequally, and reaches contained differing numbers of sites to compensate for differences in reach length (McClelland et al., 2006). Pooling data into reaches allows comparisons between reaches despite sampling inequality. Lastly, environmental data were not available for each fish survey site. The reach level approach averages out environmental differences among reaches, permits comparison between reaches despite differing numbers of sampling sites within each, and accommodated inclusion of the available environmental data.

Water Quality

With the exception of secchi visibility depth taken by LTEF, water quality parameters were obtained from the Metropolitan Water Reclamation District of Greater Chicago (MWRD). Since 1983, MWRD has collected water quality data several times annually at 49 fixed-site locations distributed within the four focal reaches as part of their Illinois Water Waterway monitoring program. Collection sites are generally located in narrow areas of the channel with plug-flow characteristics assumed to be cross-sectionally well mixed. Field measurements of dissolved oxygen, conductivity, temperature, and turbidity are taken along with a single water grab sample from the center of the channel a meter below the water surface. The water sample is preserved for United States Environmental Protection Agency approved laboratory measurements of nutrients, metals, biological oxygen demand, and suspended solids (MWRD, 2010).

Each MWRD site was associated with the nearest LTEF site, with preference given to upstream locations. In two instances, one MWRD site was used for two separate, but proximal LTEF sites. This resulted in Starved Rock having one MWRD site, and Peoria having seven, as opposed to two and eight LTEF sites respectively. Distances between LTEF and MWRD sites ranged from 0.08-4.67km with a mean of 1.96km. Following site association, the mean of MWRD variables were calculated for each reach to provide a single yearly value for each variable within each reach. In total, eleven water quality variables were obtained (Table 3.1)

Climate

The National Oceanic and Atmospheric Administration's (NOAA) National Climatic Data Center (NCDC) was accessed for regional climate data. Data were compiled from four sites with complete monthly data for the time period analyzed. Upstream to downstream, the sites were located in Chicago, Ottawa, Peru, and Peoria. Chicago climate data were used for the Dresden reach, Ottawa for the Marseilles reach, and Peru for the Starved Rock reach. Due to the relatively large size of the Peoria reach, and the locations of two climate sites at the extreme ends of the reach, climate data were averaged to better estimate climate throughout the entire reach. The monthly data were then converted to yearly climate variables for temperature and precipitation. This yielded sixteen climate variables (Table 3.1). Of these, eight corresponded to temperature, and the remaining eight to precipitation.

Functional Diversity Indices

Selection of Indices

Many FD indices have been developed (Laliberté and Legendre, 2010; Pla et al., 2012; Schleuter et al., 2010). For purposes of this study, we selected indices that met four primary requirements. First, the indices should not be correlated with species richness or each other. Second, indices need to represent different aspects of functional structure. Third, indices are able to take into account multiple traits simultaneously. Lastly, indices have the ability to be weighted by species abundance or biomass. These requirements eliminated the majority of currently available FD indices (Pla et al., 2012; Schleuter et al., 2010). Two indices met all of our requirements, functional evenness (FEve) (Villéger et al., 2008), and functional dispersion (FDis) (Laliberté and Legendre, 2010).

FEve is calculated using a minimum spanning tree to transform species distributions from multidimensional functional trait space to a single axis. Weighted branch lengths for each species are used to calculate functional trait evenness (Table 3.2). The result is an index confined between zero and one that quantifies the weighted regularity with which the functional space is filled. Lower values indicate abundances or biomass are less evenly distributed or functional distances are less regular due to species sharing less traits in common (Pool et al., 2010; Villéger et al., 2008).

FDis measures the spread of species in multidimensional trait space. It shifts the centroid toward the more abundant species and also weights distances from this centroid by relative abundance (Table 3.2) (Laliberté and Legendre, 2010; Pla et al., 2012). FDis decreases when the more abundant species in an assemblage share more traits in common. To use a foraging example, abundant species sharing the same feeding preferences would lower the value of FDis. The lower bound of FDis is 0.00, while there is no upper boundary. Additionally, as more traits are introduced to FDis the values are inflated reflecting the larger FD space.

Calculation of Functional Diversity

To calculate FD we associated fish species with the Fish Trait database (Frimpong and Angermeier, 2009). This database provides binary classification (yes/no) for a number of feeding and habitat functional traits for each species present in LTEF data. Traits were used to calculate two different aspects of functional diversity representing foraging and habitat preferences (Table 3.3). FEve and FDis indices were calculated for each reach using FDiversity software package (Casanoves et al., 2011). Both were calculated twice, weighted with either CPUEw (FEve_w, FDis_w) to reflect biomass, or CPUEn (FEve_n, FDis_n) to reflect abundance. As traits were binary categorical variables, Euclidean distance was selected for FDis calculations (Laliberté and Legendre, 2010; Pla et al., 2012).

Species richness for each reach was also calculated to draw comparisons among this traditional taxonomic diversity index and FD indices. We checked for correlation among indices by calculating Spearman rank correlation coefficients (ρ) (Legendre and Legendre, 2012). Simple ordinary least squared (OLS) regression was performed on each index using year as the predictor to assess the strength of linear temporal trends(Bingham et al., 2010). Correlation coefficients and regressions were done in InfoStat version 2011p software (Di Rienzo et al., 2011).

Modeling Indices with Environment

To assess the possible influence water quality and climate variables had on assemblages, random forest (RF) regression was applied to FEve, FDis, and species richness for each reach using package "randomForest" with R 2.14.1 software (Liaw and Wiener, 2002). RF regression is a learning algorithm that uses multiple regression trees. This algorithm does not require assumptions, permits interactions among variables, and functions well with high dimensional data (where n>>p) (Genuer et al., 2010; Genuer et al., 2008). RF provides three main outputs; 1) the percent of variation in a response variable explained by the predictors (R^2), 2) the relative importance of each predictor in explaining the variation, and 3) an approximation of the type of statistical relationship between individual predictors and the response through partial dependence plots (Breiman, 2001; Cutler et al., 2007).

RF models were run using the full suite of environmental variables as predictors on each index (Table 3.1). Three preliminary RF optimization runs were performed using a range of the number of randomly drawn variables for each split in individual regression trees (1-p/2). The lowest value for the randomly drawn variables, which came within 2% of the model that explained the largest percent of variation, was selected and ran 10 times. Due to the randomized nature of the algorithm, every run of RF generates a slightly different R^2. We selected the model for each index, which explained the greatest amount of variation (R^2) for investigation from the 10 runs, which used the optimized number of predictor variables available at each split. Failed models (explained <10% variation or negative variation) were not investigated further. Further investigation into meaningful models (explained >10% variation) included determination of the top five most important variables, and partial dependence plots of these top five variables from each model. With RF, the top five variables were identified by randomizing single variables and rerunning the model. If a variable is important to the model, when it is randomized and the model rerun, there will be an increase in the model mean squared error (MSE). Partial dependence plots attempt to average out the effects of all variables, except the one of interest. This then allows the user to infer the relationship between individual important predictors and the response. All iterations of RF utilized a forest of 5,000 regression trees to ensure stable estimates of variable importance (Liaw and Wiener, 2002).

Results

Biodiversity Indices

The three biodiversity indices varied among reaches (Figure 3.2). Species richness was similar for the three uppermost reaches (Dresden, Marseilles, Starved Rock), and higher for Peoria. Peoria also had lower FEve overall, while FDis values were similar for all four reaches (Table 3.4; Figures 3.3, 3.4). Habitat FDis was higher than foraging FDis due to the inclusion of more traits, 11 versus seven respectively. All of the indices weighted with biomass had a greater amount of variation than the abundance-weighted counterparts, as assessed by the coefficient of variation (Table 3.4).

Simple OLS regressions of each index using year as a predictor showed little linearity in temporal change for most indices (Table 3.4, Figures 3.3-3.5). Coefficients of determination ranged from 0.00 to 0.69 with a mean of 0.19, and a median of 0.10. Biomass weighted FDis ($FDis_w$) tended to have the strongest linear trends, having only one $R^2=0$ (Foraging for Peoria) and an overall mean of 0.47. FEve overall had the weak linear trends (mean $R^2 = 0.07$), and no discernable temporal patterns. Overall, $FDis_w$ increased in all reaches examined for both foraging and habitat, with the one exception of Peoria Reach foraging, which appeared stable or slightly decreasing. Abundance weighted FDis ($FDis_n$) showed little change over time in all reaches for both foraging and habitat.

The observed changes in FD appeared to be largely driven by centrarchid and cyprinid species. The members of centrarchidae contributing most strongly to abundance and biomass changes were bluegill (*Lepomis macrochirus*), green sunfish (*L. cyanellus*), and largemouth bass (*Micropterus salmoides*). These three species increased temporally in abundance and biomass, especially in the three uppermost reaches (Dresden, Marseilles, and Starved Rock). Throughout all four reaches common carp (*Cyprinus carpio*) decreased during the time period analyzed. Additional variability in biomass weighted FD was likely introduced by less often-captured large bodied benthic catfish (*Ictaluridae*) and sucker species (*Catostomidae*). Abundance weighted FD was more likely to be affected by stochastic changes in small abundant minnow (*Cyprinidae*) and shad (*Dorosoma cepedianum, D. petenense*) species.

Based upon Spearman rank correlation, only four index variations were strongly monotonically correlated with one another, $-0.5 > \rho > 0.5$ (Table 3.5). Three of the four habitat indices were strongly positively correlated with their foraging index counterpart, and species richness with habitat $FEve_w$. In every instance, species richness had a positive relationship with FDis and negative relationship with FEve.

Biodiversity-Environment Random Forest Models

In total 36 RF models were run ([Species Richness × 4 reaches] + [2 FD indices × 2 measures of density × 2 sets of traits × 4 reaches]), representing the nine diversity index variations for each of four reaches. The majority of models, 25, were unable to explain the temporal variation in diversity indices with the water quality and climate predictors. The remaining 11 models explained between 10.9% and 64.7% of the observed variation (Table 3.6). Environmental variables were able to account for variation in 1of 4 species richness model, 7 of 16 foraging FD models, and 3 of 16 habitat FD models. Of the 10 meaningful FD models, eight modeled FDis and two FEve.

Eight biomass-weighted and two abundance-weighted FD indices were meaningful (Table 3.6). Foraging based $FDis_w$ was the only biodiversity index meaningful for all four reaches, and performance decreased downstream, with an R^2 of 64.7% at Dresden, 48.0% at Marseilles, 10.9% at Starved Rock, and 15.8% at Peoria. Habitat-based $FDis_w$ were meaningfully fit for Dresden, Marseilles, and Peoria reaches, and did not display a similar downstream decrease in model performance.

Predictor variable importance varied based upon Reach, index, and weighting applied (Table 3.7, Appendix A). Of the 55 top variables (11 models × top 5 variables), 44 were water quality and 11 climate. The most important predictor variables across models include ammonia, phosphorus, phenols, and secchi visibility depth with eight, seven, seven, and six occurrences in the 11 models respectively. Peoria contained an unequal share of the climate variables with six of 10 total occurrences. Only one climate variable appeared within the top five of more than a single model, TMIN, which occurred in Dresden $FDis_w$ for both foraging and habitat.

Responses of FDis and FEve to predictors had consistent relationships with respect to the index used, except for iron, which had a positive relationship with $FDis_n$ in Dresden and negative relationship with $FDis_w$ in Peoria (Table 3.7, Appendix A). For example, ammonia was selected among the top five most important variables in seven FDis models with consistent negative effects, phosphorus in six with positive effects, phenols in 5 models with negative effects, and secchi in 4 models with positive effects. Only mercury occurred within the top five variables of both FEve models (Dresden foraging $FEve_n$, Peoria foraging $FEve_w$), and it had a positive relationship with FEve in both.

Though response predictor relationships were consistent with respect to index, there tended to be inverse relationships between FDis and FEve within reaches. Partial dependence plots for Dresden foraging $FDis_n$ showed a positive relationship with phosphorus, and negative relationship with phenols and ammonia, while foraging $FEve_n$ model showed the reverse. The same reverse relationship did not occur between foraging $FDis_w$ and $FEve_w$ in Peoria.

Discussion

The importance of functional diversity (FD) for maintaining ecosystem functions and services are increasingly recognized (Cadotte et al., 2011; Flynn et al., 2011). How FD has been restored in impaired ecosystems needs to be evaluated adequately. Our findings in this study offer insights into how functional evenness (FEve) and dispersion (FDis) responded to water quality improvements and climate variability. FDis increased over time in most reaches studied (Figure 3.3). FEve was generally much more stochastic, and showed some decrease in Dresden Reach (Figure 3.4). These trends were stronger when biomass was used to weight species than when abundance was used. Interestingly species richness steadily increased in the Peoria Reach, but FD changed little (Figure 3.5).

Previous investigations into fish FD have not included the range of water quality variables we have used. Lack of water quality data combined with a focus broad spatial scales, is a likely reason why those studies found climate to be the dominant factor driving fish FD (Mason et al., 2008; Pool et al., 2010; Schleuter et al., 2012). In the IRW, water quality clearly was the dominant factor driving the long-term fish FD (1983-2010). This is likely due to previously reported improvements in water quality affecting the structure of IRW fish assemblages (Chapter 2; Koel and Sparks, 2002; Lerczak et al., 1994; Pegg and McClelland, 2004; Sparks and Lerczak, 1993; Theriot and Tzoumis, 2005). However, unlike the previous studies, we were able to find direct statistical links between water quality and fish structure.

The negative correlation between FDis and FEve and their different responses to predictor variables to have opposing effects on these indices was interesting. A high FDis value indicates a broader breadth of species function. However, similar to species evenness, what a high FEve value means biologically is not completely clear. Negative relationships between species richness and species evenness have also been reported (Wilsey and Stirling, 2007; Wilsey et al., 2005).

Ammonia was the most important variable across all FDis models. This finding supports previous work by Sparks and others (1993; 1981), which found ammonia from sewage to be a dominant factor affecting the ecology of the upper IRW. The responses of

FDis to other important water quality variables also generally have biologically expected relationships. Dissolved oxygen, pH, and secchi depth have a positive relationship with FDis. While phenolic compounds negatively affect FDis. These results imply that FDis increased when less tolerant fishes reestablished or increased populations.

Only two FEve models were meaningful. The lack of meaningful FEve models indicates weak relationships with water quality and climate variables. One of the top predictors for foraging FEve_n in Dresden and foraging FEve_w in Peoria was mercury. In both cases, the fish assemblage became less even and more diverse as mercury decreased. This is biologically expected, as the toxicity of mercury to freshwater fishes and invertebrates at levels observed in our data is well documented (Boening, 2000). In line with FDis, reductions in ammonia concentrations increased diversity in the Dresden FEve model. Sparks and Sandusky (1983) did find toxic levels of ammonia in the Upper IRW sediment at the beginning of the time period we investigated. We surmise the decreased ammonia had a positive effect either directly on fish biology, or on invertebrates, which increased the forage base for fishes.

The finding that the majority (70%) of meaningful models were foraging-based supports our hypothesis that foraging FD was more affected by the change in water quality and climate. However, the three habitat FDis_w models also performed well. We attribute this result to increased use of available habitats due to improved water quality supporting fish species with diverse habitat preferences, and not to increases in habitat diversity *per se*. However, since we lacked habitat data in our study, this is only conjecture based upon the fixed site methodology and the absence of change to flow regime modifications.

Because 80% of meaningful FD models are biomass-based, biomass is clearly more effective for capturing FD response to environment. To our knowledge, this is the first use of biomass in calculation of FD indices for fish assemblages. This is not surprising due to the technical difficulties inherent when assessing biomass of individual species. However, given the extreme size difference present among and within fish species, and our results showing the efficacy of using biomass, biomass be used whenever possible. Using growth curve estimations or assigning small fish a predetermined minimum value, should be viewed as preferable to the possible loss of biologically meaningful environmental relationships (Chen et al., 1992).

When characterizing an assemblage's response to their environment, it is important to examine multiple aspects of structure. We have demonstrated functional diversity has the potential to depict responses, which may go undetected using the traditional taxonomic diversity index of species richness. In our study, we found strong relationships between environmental variables and FD indices, but not species richness. These results are significant, as Cadotte et al. (2011) pointed out the lack of studies assessing FD trends in relation to abiotic factors. Uses of FD indices are still fairly new, and establishing biologically meaningful relationships with the environment are necessary to advance our understanding of how assemblages are structured.

References

Bellard, C, Bertelsmeier, C, Leadley, P, Thuiller, W and Courchamp, F. 2012. Impacts of climate change on the future of biodiversity. *Ecology Letters* **15**: 365-377.

Bingham, NH, Bingham, N and Fry, JM. 2010. *Regression: Linear Models in Statistics*. Springer, London, UK.

Boening, DW. 2000. Ecological effects, transport, and fate of mercury: a general review. *Chemosphere* **40**: 1335-1351.

Breiman, L. 2001. Random Forests. *Machine Learning* **45**: 5-32.

Bunn, SE and Arthington, AH. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management* **30**: 492-507.

Butts, TA, Evans, RL and Lin, S. 1975. Water quality features of the Upper Illinois Waterway. Illinois State Water Survey, Champaign, IL.

Cadotte, MW, Carscadden, K and Mirotchnick, N. 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology* **48**: 1079-1087.

Cardinale, BJ, Wright, JP, Cadotte, MW, Carroll, IT, Hector, A, Srivastava, DS, Loreau, M and Weis, JJ. 2007. Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proceedings of the National Academy of Sciences* **104**: 18123-18128.

Carey, MP and Wahl, DH. 2011. Determining the mechanism by which fish diversity influences production. *Oecologia* **167**: 189-198.

Casanoves, F, Pla, L, Di Rienzo, JA and Díaz, S. 2011. FDiversity: a software package for the integrated analysis of functional diversity. *Methods in Ecology and Evolution* **2**: 233-237.

Chen, Y, Jackson, D and Harvey, H. 1992. A comparison of von Bertalanffy and polynomial functions in modelling fish growth data. *Canadian Journal of Fisheries and Aquatic Sciences* **49**: 1228-1235.

Cohen, JE, Jonsson, T and Carpenter, SR. 2003. Ecological community description using the food web, species abundance, and body size. *Proceedings of the National Academy of Sciences* **100**: 1781-1786.

Cutler, DR, Edwards Jr, TC, Beard, KH, Cutler, A, Hess, KT, Gibson, J and Lawler, JJ. 2007. Random forests for classification in ecology. *Ecology* **88**: 2783-2792.

Dauer, D, Luckenbach, M and Rodi Jr, A. 1993. Abundance biomass comparison (ABC method): effects of an estuarine gradient, anoxic/hypoxic events and contaminated sediments. *Marine Biology* **116**: 507-518.

Di Rienzo, JA, Casanoves, F, Balzarini, MG, Gonzalez, L, Tablada, M and Robledo, CW. 2011. InfoStat 2011p.

Dudgeon, D, Arthington, AH, Gessner, MO, Kawabata, ZI, Knowler, DJ, Lévêque, C, Naiman, RJ, Prieur-Richard, AH, Soto, D and Stiassny, ML. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* **81**: 163-182.

Ernst, R, Linsenmair, K and Rodel, M. 2006. Diversity erosion beyond the species level: dramatic loss of functional diversity after selective logging in two tropical amphibian communities. *Biological Conservation* **133**: 143-155.

Flynn, DF, Mirotchnick, N, Jain, M, Palmer, MI and Naeem, S. 2011. Functional and phylogenetic diversity as predictors of biodiversity-ecosystem-function relationships. *Ecology* **92**: 1573-1581.

Flynn, DFB, Gogol-Prokurat, M, Nogeire, T, Molinari, N, Richers, BT, Lin, BB, Simpson, N, Mayfield, MM and DeClerck, F. 2009. Loss of functional diversity under land use intensification across multiple taxa. *Ecology Letters* **12**: 22-33.

Frimpong, EA and Angermeier, PL. 2009. Fish Traits: a database of ecological and life-history traits of freshwater fishes of the United States. *Fisheries* **34**: 487-495.

Genuer, R, Poggi, J-M and Tuleau-Malot, C. 2010. Variable selection using random forests. *Pattern Recognition Letters* **31**: 2225-2236.

Genuer, R, Poggi, J-M and Tuleau, C. 2008. Random Forests: some methodological insights. *Research report INRIA Saclay* **RR-6729**: 1-32.

Guégan, J-F, Lek, S and Oberdorff, T. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* **391**: 382-384.

Koel, TM and Sparks, RE. 2002. Historical patterns of river stage and fish communities as criteria for operations of dams on the Illinois River. *River Research and Applications* **18**: 3-19.

Laliberté, E and Legendre, P. 2010. A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **91**: 299-305.

Legendre, P and Legendre, L. 2012. *Numerical Ecology*. Elsevier.

Lerczak, TV, Sparks, RE and Blodgett, DK. 1994. The long-term Illinois River fish population monitoring program project F-101-R. Illinois Natural History Survey, Champaign, IL.

Lian, Y, You, J, Sparks, R and Demissie, M. 2012. Impact of human activities to hydrologic alterations on the Illinois River. *Journal of Hydrologic Engineering* **17**: 537-546.

Liaw, A and Wiener, M. 2002. Classification and regression by randomForest. *R News* **2/3**: 18-22.

Loreau, M. 2010. Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**: 49-60.

Mason, NW, Irz, P, Lanoiselée, C, Mouillot, D and Argillier, C. 2008. Evidence that niche specialization explains species–energy relationships in lake fish communities. *Journal of Animal Ecology* **77**: 285-296.

McClelland, MA, Pegg, MA and Spier, TW. 2006. Longitudinal patterns of the Illinois Waterway fish community. *Journal of Freshwater Ecology* **21**: 91-99.

McClelland, MA and Sass, GG. 2012. Assessing fish collections from random and fixed site sampling methods on the Illinois River. *Journal of Freshwater Ecology* **27**: 325-333.

McClelland, MA, Sass, GG, Cook, TR, Irons, KS, Michaels, NN, O'Hara, TM and Smith, CS. 2012. The long-term Illinois River fish population monitoring program. *Fisheries* **37**: 340-350.

McGill, BJ, Enquist, BJ, Weiher, E and Westoby, M. 2006. Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution* **21**: 178-185.

Meybeck, M. 2003. Global analysis of river systems: from Earth system controls to Anthropocene syndromes. *Philosophical Transactions of the Royal Society of London: Biological Sciences* **358**: 1935-1955.

Meybeck, M. 2004. The global change of continental aquatic systems: dominant impacts of human activities. *Water Science and Technology* **49**: 73.

MWRD. 2010. Illinois Waterway monitoring project quality assurance project plan. Monitoring and Research Department, Chicago, IL.

Naeem, S, Knops, JM, Tilman, D, Howe, KM, Kennedy, T and Gale, S. 2000. Plant diversity increases resistance to invasion in the absence of covarying extrinsic factors. *Oikos* **91**: 97-108.

Pease, AA, González-Díaz, AA, Rodiles-Hernández, R and Winemiller, KO. 2012. Functional diversity and trait–environment relationships of stream fish assemblages in a large tropical catchment. *Freshwater Biology* **57**: 1060-1075.

Pegg, MA and McClelland, MA. 2004. Spatial and temporal patterns in fish communities along the Illinois River. *Ecology of Freshwater Fish* **13**: 125-135.

Pla, LE, Casanoves, F and Di Rienzo, J. 2012. *Quantifying Functional Biodiversity*. Springer.

Pool, TK, Olden, JD, Whittier, JB and Paukert, CP. 2010. Environmental drivers of fish functional diversity and composition in the Lower Colorado River Basin. *Canadian Journal of Fisheries and Aquatic Sciences* **67**: 1791-1807.

Ricciardi, A and Rasmussen, JB. 1999. Extinction Rates of North American Freshwater Fauna. *Conservation Biology* **13**: 1220-1222.

Sala, OE, III, FSC, Armesto, JJ, Berlow, E, Bloomfield, J, Dirzo, R, Huber-Sanwald, E, Huenneke, LF, Jackson, RB, Kinzig, A, Leemans, R, Lodge, DM, Mooney, HA, Oesterheld, Mn, Poff, NL, Sykes, MT, Walker, BH, Walker, M and Wall, DH. 2000. Global Biodiversity Scenarios for the Year 2100. *Science* **287**: 1770-1774.

Schleuter, D, Daufresne, M, Massol, F and Argillier, C. 2010. A user's guide to functional diversity indices. *Ecological Monographs* **80**: 469-484.

Schleuter, D, Daufresne, M, Veslot, J, Mason, NW, Lanoiselée, C, Brosse, S, Beauchard, O and Argillier, C. 2012. Geographic isolation and climate govern the functional diversity of native fish communities in European drainage basins. *Global Ecology and Biogeography* **21**: 1083-1095.

Sparks, RE. 2010. Forty years of science and management on the Upper Mississippi River: an analysis of the past and a view of the future. *Hydrobiologia* **640**: 3-15.

Sparks, RE and Lerczak, TV. 1993. Recent trends in the Illinois River indicated by fish populations. Illinois Natural History Survey, Champaign, IL.

Sparks, RE and Sandusky, MJ. 1981. Identification of factors responsible for decreased production of fish food organisms in the Illinois and Mississippi Rivers. Illinois Natural History Survey, Champaign, IL.

Sparks, RE and Sandusky, MJ. 1983. Identification of the water quality factors which prevent fingernail clams from recolonizing the Illinois River phase: phase III. Illinois Natural History Survey, Champaign, IL.

Theriot, C and Tzoumis, K. 2005. Deep tunnels and fried fish: tracing the legacy of human interventions on the Chicago River. *Golden Gate University of Law Review* **35**: 377.

Tilman, D. 2001. Functional diversity. *Encyclopedia of Biodiversity* **3**: 109-120.

Villéger, S, Mason, NW and Mouillot, D. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* **89**: 2290-2301.

Villéger, S, Miranda, JR, Hernández, DF and Mouillot, D. 2010. Contrasting changes in taxonomic vs. functional diversity of tropical fish communities after habitat degradation. *Ecological Applications* **20**: 1512-1522.

Wilsey, B and Stirling, G. 2007. Species richness and evenness respond in a different manner to propagule density in developing prairie microcosm communities. *Plant Ecology* **190**: 259-273.

Wilsey, BJ, Chalcraft, DR, Bowles, CM and Willig, MR. 2005. Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity. *Ecology* **86**: 1178-1184.

Figures and Tables: Chapter 3

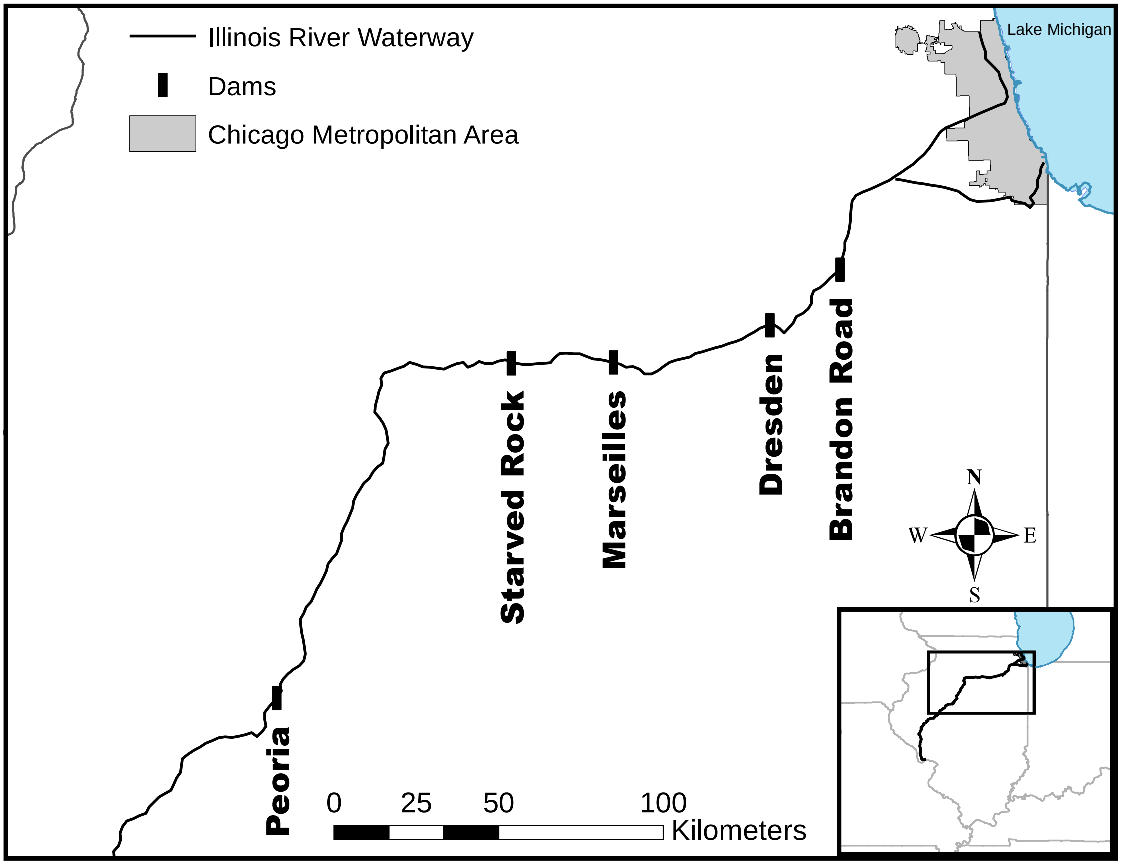


Figure 3.1: Map of our study area. The Upper Illinois River Waterway is divided into four reaches defined by lock and dam complexes. Reaches are named for pools formed by the dam located at their downstream extent.

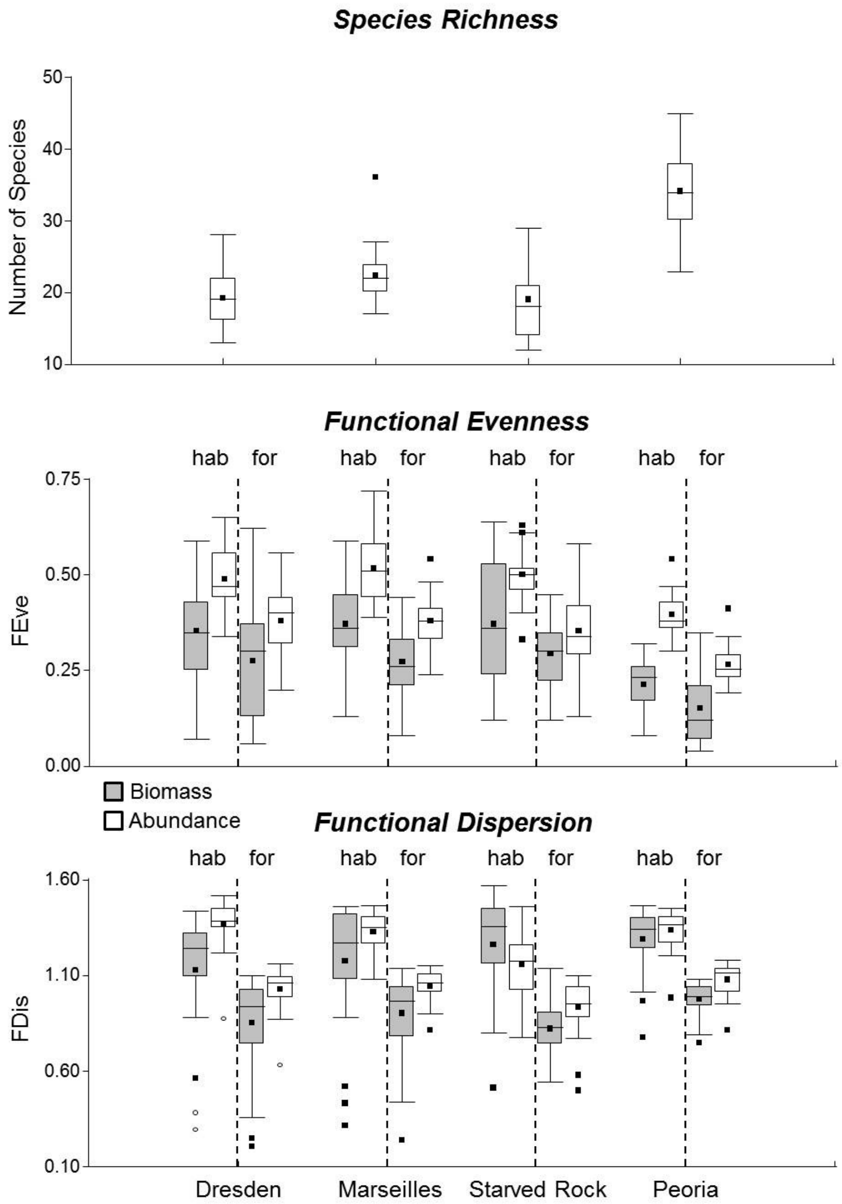


Figure 3.2; Mean and ranges of species richness and functional indices between 1983 and 2010 in each of the 4 reaches of the Illinois River Waterway included in this study. For functional evenness (FEve) and functional dispersion (FDis) the left side of dotted line represents habitat-based (hab), while the right side represents foraging-based (for). Dark shading denotes biomass-weighted and white denotes abundance-weighted.

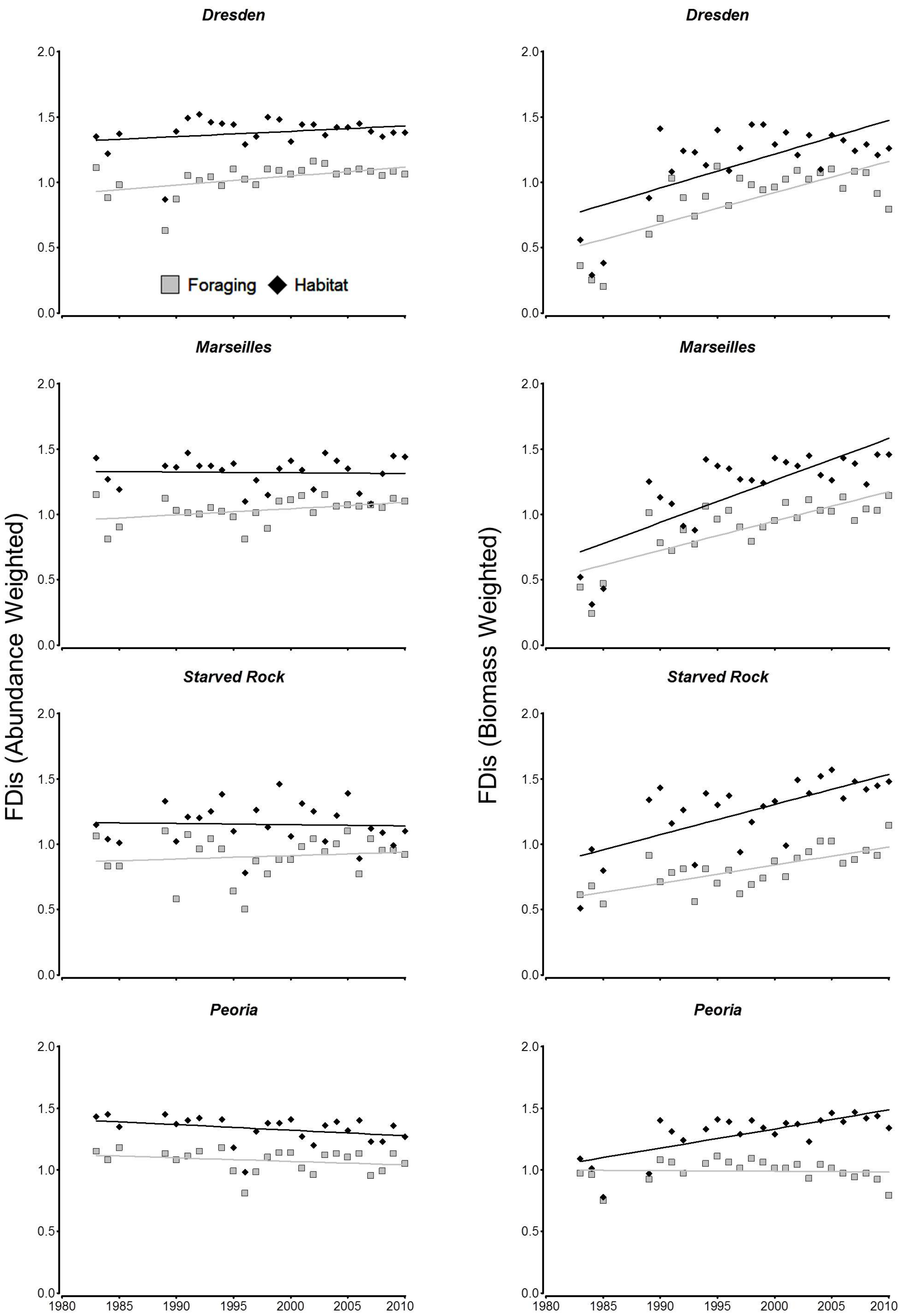


Figure 3.3: Temporal trends in functional dispersion (FDis) for habitat and foraging traits weighted by biomass or abundance in each of the four reaches of the Illinois River Waterway studied.

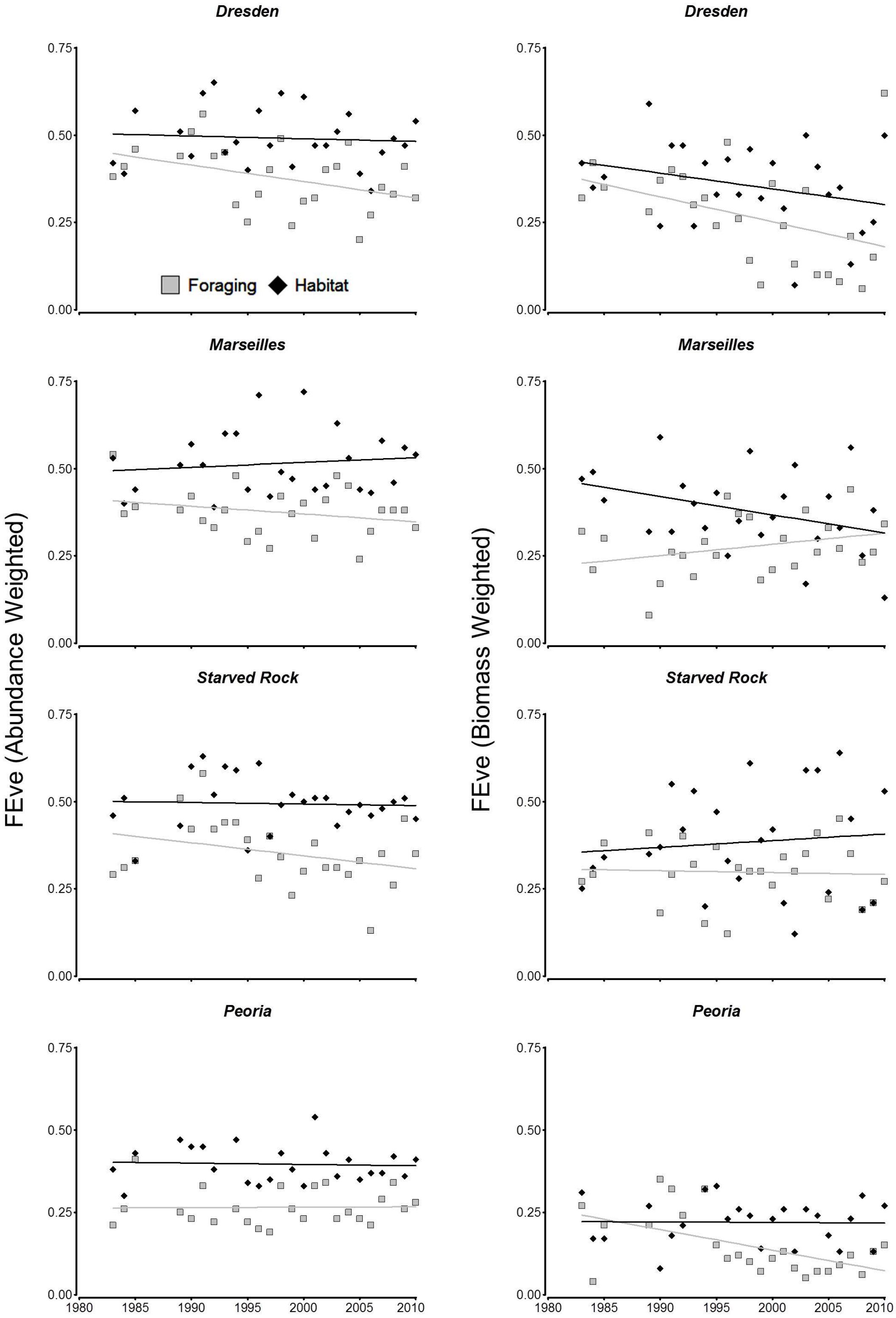


Figure 3.4: Temporal trends in functional evenness (FEve) for habitat and foraging traits weighted by biomass or abundance in each of the four reaches of the Illinois River Waterway studied.

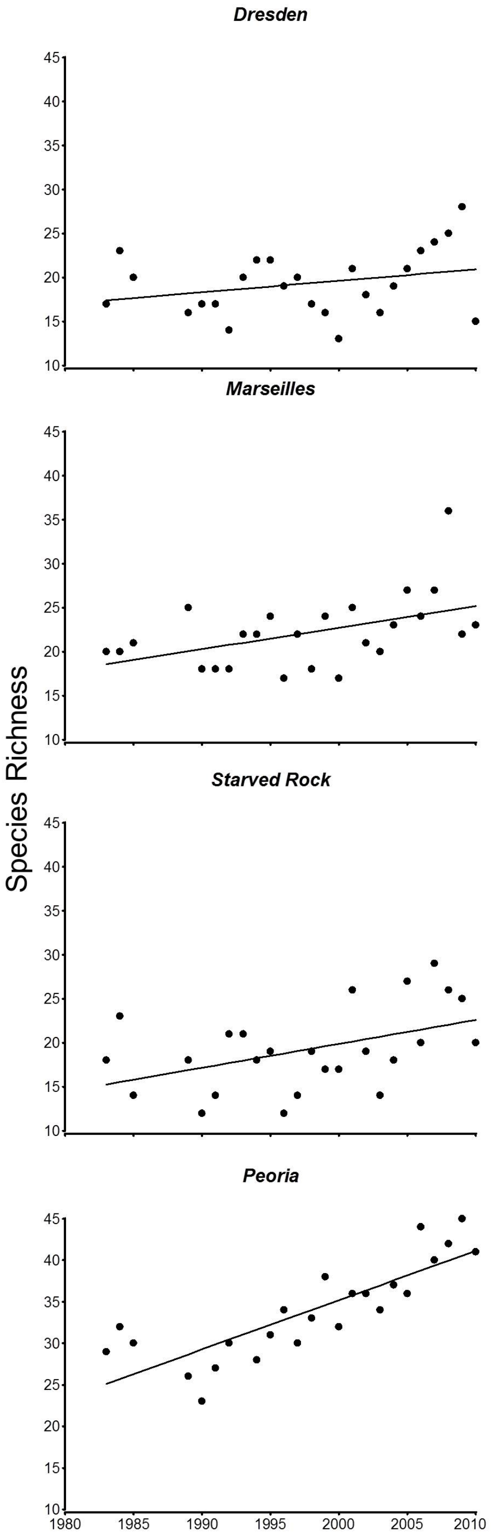


Figure 3.5: Long-term trends in species richness in each of the four reaches of the Illinois River Waterway studied.

Table 3.1: Environmental variables used in random forest regression models.

Variable	Definition	Median	Range	Units
Water Quality				
Secchi	Secchi Visibility Depth	30.7	10.6 - 93	cm
FecCol	Fecal Coliform	118.6	10.8 - 3093.3	cfu/dl
DO	Dissolved Oxygen	7.96	5.09 - 11.18	mg/l
BOD5	5-Day Biological Oxygen Demand	3.31	0 - 9.5	mg/l
TSS	Total Suspended Solids	28.7	11.2 - 102.8	mg/l
NH3	Ammonia	0.14	0 - 2.4	mg/l
P	Phosphorus	0.55	0.29 - 1.26	mg/l
pH	pH	7.7	6.7 - 8.7	
Phen	Phenolic Compounds	5.36	0.94 - 30.40	μg/l
Fe	Iron	0.6	0.2 - 8.9	mg/l
Mn	Manganese	0.05	0.03 - 0.41	mg/l
Hg	Mercury	0.07	0 - 0.25	mg/l
Climate				
DP01	Days with Precipitation ≥ 0.1"	70	50 – 93	Days
DP05	Days with Precipitation ≥ 0.5"	24	12 – 37	Days
DP10	Days with Precipitation ≥ 1.0"	8	0 – 15	Days
DT00	Days with Min Temp < 0 F	9	1 – 28	Days
DT90	Days with Max Temp ≥ 90 F	17	1 – 50	Days
DX32	Days with Max Temp < 32 F	45	15 – 71	Days
PMAX	Yearly Mean Monthly Max Rainfall	17.3	10.3 - 36.2	cm
PTOT	Total Yearly Precipitation	94.0	56.1 - 137.3	cm
PAVG	Yearly Mean Monthly Mean Rainfall	7.8	4.7 - 11.4	cm
PMIN	Yearly Mean Monthly Min Rainfall	1.8	0 - 4.2	cm
TSNW	Total Yearly Snowfall	67	14 – 173	cm
TMIN	Yearly Mean of Extreme Min Monthly Temps	-3.8	-7 - -0.2	°C
TMAX	Yearly Mean of Extreme Max Monthly Temps	25.4	23.2 - 27.5	°C
TMNA	Yearly Mean of Mean Monthly Min Temps	-10.9	-15.7 - -4	°C
TMXA	Yearly Mean of Mean Monthly Max Temps	15.8	13.9 - 18.3	°C
TAVG	Yearly Mean of Mean Monthly Temps	10.4	8.0 - 12.2	°C

Table 3.2: Two functional diversity measures, functional evenness (FEve) and functional dispersion (FDis)

Calculation	**Equation**	**Definition of terms**
Functional Evenness (FEve)		
Weighted Evenness	$EW_l = \frac{dist(i,j)}{w_i + w_j}$	EW_l = weighted evenness of branch *l* $dist(i,j)$ = euclidean distance between species *i* and *j* in branch *l* w_i = relative abundance or biomass of species *i* w_j = relative abundance or biomass of species *j*
Partial Weighted Evenness	$PEW_l = \frac{EW_l}{\sum_{l=1}^{s-1} EW_l}$	PEW_l = partial weighted evenness of branch *l* EW_l = weighted evenness of branch *l*
Functional Evenness	$FEve = \frac{\sum_{l-1}^{s-1} \min\left(PEW_l, \frac{1}{S-1}\right) - \frac{1}{S-1}}{1 - \frac{1}{S-1}}$	$FEve$ = functional evenness PEW_l = partial weighted evenness of branch *l* S = number of species
Functional Dispersion (FDis)		
Weighted Centroid	$c = [c_i] = \frac{\sum a_j x_{ij}}{\sum a_j}$	c_i = weighted centroid in *i*-dimensional space a_j = abundance or biomass of species *j* x_{ij} = attribute of species *j* for trait *i*
Functional Dispersion	$FDis = \frac{\sum a_j z_j}{\sum a_j}$	$FDis$ = functional dispersion a_j = abundance or biomass of species *j* z_j = distance of species *j* to weighted centroid

Table 3.3: Fish functional traits used in calculation of FEve and FDis for habitat and foraging from the FishTrait database (Frimpong and Angermeier, 2009).

Habitat Traits	Foraging Traits
1. Muck substrate	1. Benthic feeding
2. Clay or silt substrate	2. Surface or water column feeding
3. Sand substrate	3. Algae or phytoplankton consumption
4. Gravel substrate	4. Macrophyte and vascular plant consumption
5. Cobble or pebble substrate	5. Detritus or unidentifiable vegetative matter consumption
6. Boulder substrate	6. Larger fish, crayfish, crabs, frog consumption
7. Bedrock substrate	7. Eggs consumption
8. Aquatic vegetation substrate	
9. Organic debris or detrital substrate	
10. Large woody debris	
11. Open water	

Table 3.4: Summary of diversity indices by reach (S.D. = standard deviation, CV = coefficient of variation). The R^2 value is the result of simple linear regression with year as the single predictor variable.

Reach	Index	n	Mean	Range	S.D.	CV	R^2
Species Richness							
Dresden	S	25	19.32	13 -28	3.66	18.9	0.08
Marseilles	S	25	22.16	17 -36	4.11	18.6	0.23
Starved Rock	S	25	19.24	12 -29	4.71	24.5	0.21
Peoria	S	24	33.92	23 -45	5.78	17.0	0.69
Foraging							
Dresden	FDis_n	25	1.03	0.63 -1.16	0.11	10.6	0.25
	FDis_w	25	0.86	0.20 -1.12	0.26	30.2	0.53
	FEve_n	25	0.38	0.20 -0.56	0.09	24.1	0.17
	FEve_w	25	0.27	0.06 -0.62	0.14	53.0	0.16
Marseilles	FDis_n	25	1.03	0.81 -1.15	0.10	9.2	0.16
	FDis_w	25	0.90	0.24 -1.14	0.23	25.3	0.62
	FEve_n	25	0.38	0.24 -0.54	0.07	18.4	0.07
	FEve_w	25	0.28	0.08 -0.44	0.08	30.3	0.09
Starved Rock	FDis_n	25	0.91	0.50 -1.10	0.16	17.4	0.10
	FDis_w	25	0.81	0.54 -1.14	0.15	18.7	0.54
	FEve_n	25	0.35	0.13 -0.58	0.09	26.5	0.10
	FEve_w	25	0.30	0.12 -0.45	0.09	28.7	0.00
Peoria	FDis_n	24	1.07	0.81 -1.18	0.09	8.4	0.07
	FDis_w	24	0.99	0.75 -1.11	0.09	8.8	0.00
	FEve_n	24	0.27	0.19 -0.41	0.06	21.0	0.00
	FEve_w	24	0.15	0.04 -0.35	0.09	61.7	0.30
Habitat							
Dresden	FDis_n	25	1.38	0.87 -1.52	0.13	9.2	0.07
	FDis_w	25	1.15	0.29 -1.44	0.31	27.0	0.44
	FEve_n	25	0.49	0.34 -0.65	0.08	16.8	0.01
	FEve_w	25	0.36	0.07 -0.59	0.12	33.9	0.09
Marseilles	FDis_n	25	1.32	1.08 -1.47	0.12	8.8	0.01
	FDis_w	25	1.18	0.31 -1.46	0.33	27.7	0.61
	FEve_n	25	0.51	0.39 -0.72	0.09	17.6	0.01
	FEve_w	25	0.38	0.13 -0.59	0.12	30.4	0.13
Starved Rock	FDis_n	25	1.15	0.78 -1.46	0.16	14.0	0.00
	FDis_w	25	1.25	0.51 -1.57	0.27	21.3	0.48
	FEve_n	25	0.49	0.33 -0.63	0.07	15.1	0.00
	FEve_w	25	0.38	0.12 -0.64	0.15	39.4	0.01
Peoria	FDis_n	24	1.33	0.98 -1.45	0.11	8.2	0.11
	FDis_w	24	1.30	0.78 -1.47	0.17	13.2	0.53
	FEve_n	24	0.40	0.30 -0.54	0.06	14.1	0.00
	FEve_w	24	0.22	0.08 -0.33	0.07	30.9	0.00

Table 3.5: Spearman rank correlations among multiple diversity measures. (* indicates a strong relationship between indices > 0.50 or <-0.50)

			Foraging				***Habitat***			
		S	**FDis$_n$**	**FDis$_w$**	**FEve$_n$**	**FEve$_w$**	**FDis$_n$**	**FDis$_w$**	**FEve$_n$**	**FEve$_w$**
	S	1.00								
Foraging	**FDis$_n$**	0.35	1.00							
	FDis$_w$	0.39	0.43	1.00						
	FEve$_n$	-0.43	-0.12	-0.29	1.00					
	FEve$_w$	-0.48	-0.24	-0.29	0.30	1.00				
Habitat	**FDis$_n$**	0.05	0.63*	0.27	0.01	-0.18	1.00			
	FDis$_w$	0.19	0.08	0.55*	-0.22	-0.17	-0.05	1.00		
	FEve$_n$	-0.49	-0.10	-0.11	0.56*	0.34	0.04	0.00	1.00	
	FEve$_w$	-0.54*	-0.32	-0.35	0.27	0.46	-0.24	-0.21	0.30	1.00

Table 3.6: Summary of random forest regression models with $R^2 > 10\%$. The percentage within cells is the amount of variation explained in a given response variable by the water quality and climate predictors.

		Foraging FD				***Habitat FD***			
Reach	**S**	**$FDis_n$**	**$FDis_w$**	**$FEve_n$**	**$FEve_w$**	**$FDis_n$**	**$FDis_w$**	**$FEve_n$**	**$FEve_w$**
Dresden	-	25.5%	64.7%	16.6%	-	-	47.2%	-	-
Marseilles	-	-	48.0%	-	-	-	55.6%	-	-
Starved Rock	-	-	10.9%	-	-	-	-	-	-
Peoria	44.13%	-	15.8%	-	18.0%	-	39.7%	-	-

Table 3.7: Summary of random forest regression models with list of top five predictor variables for each model. The importance of each variable is quantified by the increase in model mean squared error when the variable is randomized. Relationships between each predictor and the response were garnered from partial dependence plots. *(NL) = non-linear relationship

Reach	**Traits**	**Index**	**%Variation Explained**	**Top 5 Predictors**	**%Increase in MSE**	**Relationship**
Dresden	***Foraging***	**FDis$_n$**	25.46%	P	27.45%	Positive
				Fe	13.24%	Positive
				Phen	12.89%	Negative
				NH3	12.66%	Negative
				DO	10.81%	Positive
		FDis$_w$	64.66%	P	25.74%	Positive
				Secchi	25.40%	Positive
				NH3	20.82%	Negative
				TMIN	17.98%	Positive (NL)
				Phen	15.89%	Negative
		FEve$_n$	16.58%	Hg	29.91%	Positive
				NH3	13.18%	Positive
				P	13.08%	Negative
				TSNW	9.58%	Negative
				Phen	8.84%	Positive
	Habitat	**FDis$_w$**	47.22%	TMIN	22.53%	Positive
				NH3	18.51%	Negative
				Secchi	17.96%	Positive
				Phen	16.30%	Negative
				P	10.92%	Positive
Marseilles	***Foraging***	**FDis$_w$**	47.98%	NH3	24.69%	Negative
				pH	23.71%	Positive
				Phen	16.25%	Negative
				P	13.34%	Positive
				DP10	9.75%	Negative
	Habitat	**FDis$_w$**	55.62%	pH	37.35%	Positive
				NH3	23.93%	Negative
				P	17.88%	Positive
				Secchi	14.42%	Positive
				Hg	10.55%	Negative

Table 3.7: Continued

Reach	Traits	Index	%Variation Explained	Top 5 Predictors	%Increase in MSE	Relationship
Starved Rock	***Foraging***	**FDis$_w$**	10.93%	DO	23.33%	Positive
				Secchi	11%	Positive
				NH3	9.83%	Negative (NL)
				P	7.91%	Positive
				TMNA	7.57%	Negative
Peoria	***None***	**Species Richness**	44.13%	Fe	31.54%	Positive
				Secchi	28.56%	Positive
				DO	19.54%	Positive
				Phen	15.98%	Negative
				PMAX	10.06%	Positive
	Foraging	**FDis$_w$**	15.80%	DX32	22.37%	Negative
				Fe	19.39%	Negative
				TMAX	16.07%	Positive
				BOD	8.22%	Negative
				pH	7.82%	Positive
		FEve$_w$	18.01%	Secchi	37.31%	Negative
				Hg	15.37%	Positive
				TMXA	10.46%	Positive
				Fe	7.62%	Negative
				TAVG	7.22%	Positive
	Habitat	**FDis$_w$**	39.73%	pH	26.16%	Positive
				Phen	19.97%	Negative
				NH3	16.66%	Negative
				DT00	10.07%	Negative
				DO	9.60%	Positive

Chapter 4

Summary

Human alterations have had significant impacts on large river ecosystems (Meybeck, 2004). These impacts are well documented, and much effort, often spurred by legislation, has been put toward restoring ecosystems (Andrews and Rickard, 1980; Pegg and McClelland, 2004; Plum and Schulte-Wülwer-Leidig, 2012). Unfortunately, we still cannot accurately predict how assemblages will respond to restoration attempts. To manage ecosystems properly, we must first understand the impacts of restoration activities (e.g. improved water quality) on assemblages.

My thesis made novel utilization of previously well-analyzed long-term fish data by coupling it with detailed water quality and climate data. Both environmental factors are known to alter abundance, biomass, and diversity of fishes (Clarke and Johnston, 1999; Poff et al., 2007; Poulin, 1992; Randall and Tsui, 2002). In Chapter 2, I examined species diversity, and six proportional attributes related to native and predatory fishes. I found predatory fish responded most strongly to changes in water quality. In Chapter 3, I examined changes in foraging-based and habitat-based functional diversity in relation to water quality and climate, and found biomass based models outperformed abundance based models. In the analyses of both chapters there tended to be decreases in the strength of the assemblage-environment relationships downstream. This observation combined with stronger assemblage attribute responses to sewage effluent related water quality variables (i.e. fecal coliform) upstream, allowed me to conclude Chicago area sewage was a dominant factor structuring upstream fish assemblages over the period-analyzed (1983-2010).

Both chapters demonstrate species richness, whether calculated as either reach mean or summed, was a less robust indicator of overall assemblage response to the environmental variables tested than other response variables investigated. I contend one of the most scientifically novel findings are the stronger effects of water quality and climate variables on biomass-based functional diversity indices than ones based on abundance. This has profound implications for future functional diversity research, as it displays a lack of detection power in the more commonly used abundance-based indices and highlights the importance of measuring biomass in fish surveys.

Overall, my study indicates multiple aspects of fish assemblage structure are necessary to ascertain environmental responses, as some aspects were only weakly correlated with the abiotic variables. Combined, the results of both studies indicate a) biomass-based assemblage estimates can provide a much stronger and more biologically meaningful relationship to water quality and climate; b) a large portion of assemblage changes can be explained by improved water quality; c) the water quality variables fishes are responding to most strongly are clarity, ammonia, and dissolved oxygen; d) fish changes in reaches closer to Chicago are more strongly correlated with water quality. These results imply improvements in water quality have led to more diverse fish assemblages.

There are a few scientific and management applications of my studies. Primarily, improving water quality may be an effective means to improving fish diversity. This conclusion is partially opposed to management's focus on physical habitat in lotic systems, the "if you build it they will come" approach (Bond and Lake, 2003), and is especially pertinent to large regulated rivers where habitat and flow modifications are largely infeasible. Second, because of the differential effects of water quality variables on fish assemblages, mitigation and monitoring efforts may want to focus on specific pollutants (i.e. ammonia, suspended solids, nutrients). The main limitation of my studies lies in the inability to predict if and how the fish assemblages will continue to change with additional water quality improvements expected as a result of the Deep Tunnel Project (Scalise and Fitzpatrick, 2012). There is no reference condition for the IRW. When monitoring began, the IRW was already heavily impacted by human alterations. Moreover, there are no similar unmodified large rivers with the same species pool to obtain a reference condition on which to base predictions. Based upon my research showing Peoria Reach to be fairly stable, I would predict that upstream reaches might begin to become more like Peoria Reach. However, this prediction may not be realistic, as reaches have different geomorphology, land cover, and hydrology.

Finally, there are further needs and directions continuing Illinois River research should take. First, given the more ecologically meaningful results of biomass-based functional diversity estimates, I recommend that biomass measurements of even small fish be included in future data collections. Due to the time and effort necessary for this to occur, it may be beneficial to establish more accurate growth curves for small fish species to obtain better estimates of weight. Next, due to the strong relationship between water quality and fish assemblage characteristics, I recommend LTEF consider collecting a range of water quality data throughout the entire sampling area. This would allow future researchers to quantify the effects of water quality on fish on a larger spatial scale, and would be a great benefit to ecologists and managers seeking to improve fish diversity throughout the world. Finally, as phosphorus was the only variable, which showed impacts on fish contrary to general expectations, an investigation into the sources and forms of phosphorus in the Upper Illinois River is needed.

In summary, I evaluated the long-term changes in multiple attributes of Illinois River Waterway fish assemblages in relation to water quality improvements and interannual climate variability. I identified a set of water quality and climate variables that best explained the fish assemblage changes. The combined results of both data chapters shows the advantages of multi-agency data sharing, display the utility of using multiple aspects of assemblage structure in order to determine environmental impacts, validate the collection of long-term data, and indicate improvements in sewage handling are aiding in the recovery of Illinois River Waterway fishes.

References

Andrews, M and Rickard, D. 1980. Rehabilitation of the inner Thames estuary. *Marine Pollution Bulletin* **11**: 327-332.

Bond, NR and Lake, PS. 2003. Local habitat restoration in streams: constraints on the effectiveness of restoration for stream biota. *Ecological Management & Restoration* **4**: 193-198.

Clarke, A and Johnston, NM. 1999. Scaling of metabolic rate with body mass and temperature in teleost fish. *Journal of Animal Ecology* **68**: 893-905.

Meybeck, M. 2004. The global change of continental aquatic systems: dominant impacts of human activities. *Water Science and Technology* **49**: 73.

Pegg, MA and McClelland, MA. 2004. Spatial and temporal patterns in fish communities along the Illinois River. *Ecology of Freshwater Fish* **13**: 125-135.

Plum, N and Schulte-Wülwer-Leidig, A. 2012. From a sewer into a living river: the Rhine between Sandoz and Salmon. *Hydrobiologia*: 1-12.

Poff, NL, Olden, JD, Merritt, DM and Pepin, DM. 2007. Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences* **104**: 5732-5737.

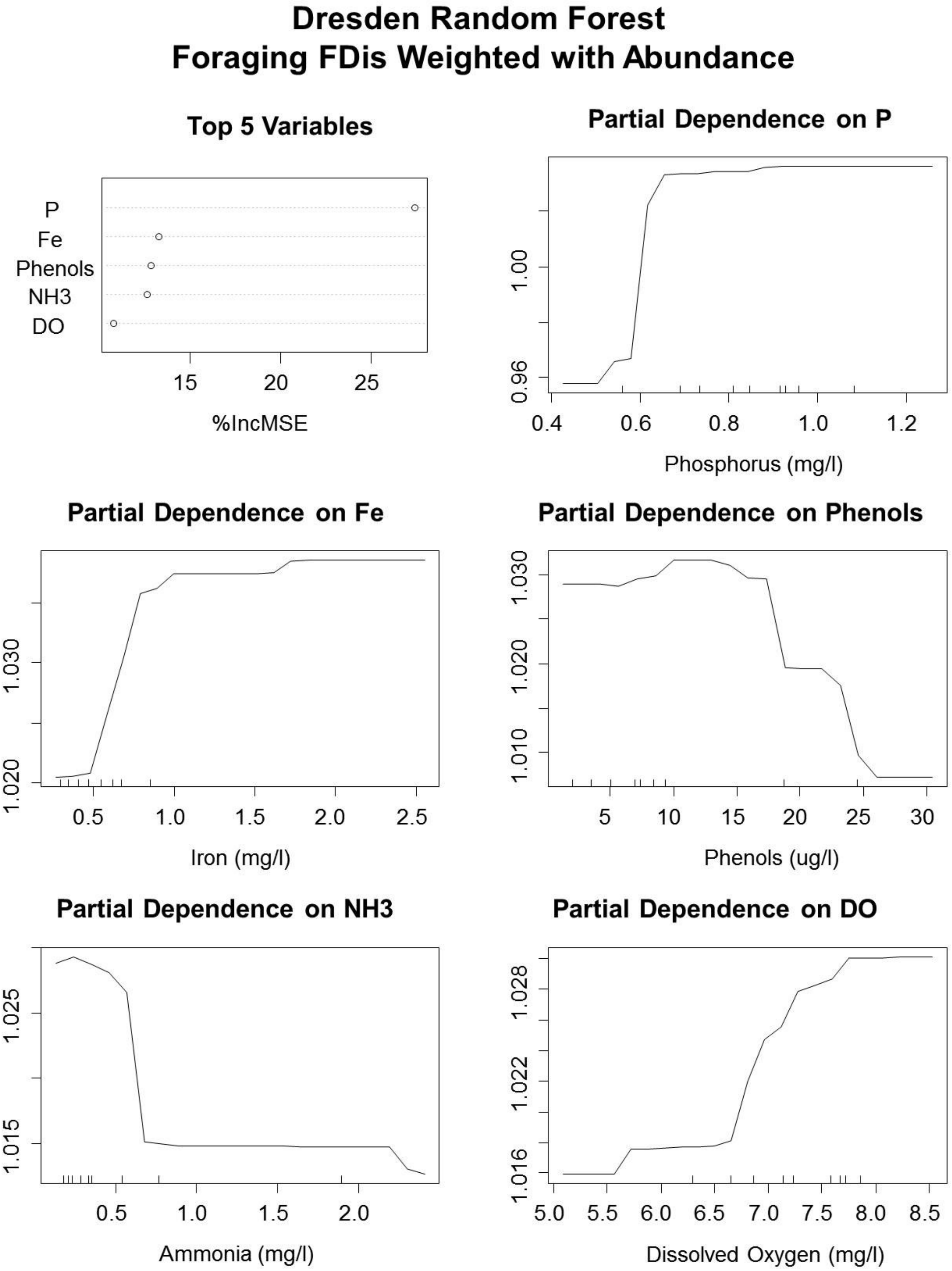
Poulin, R. 1992. Toxic pollution and parasitism in freshwater fish. *Parasitology Today* **8**: 58-61.

Randall, D and Tsui, T. 2002. Ammonia toxicity in fish. *Marine Pollution Bulletin* **45**: 17-23.

Scalise, C and Fitzpatrick, K. 2012. Chicago deep tunnel design and construction. *Structures Congress 2012.* Chicago, IL.

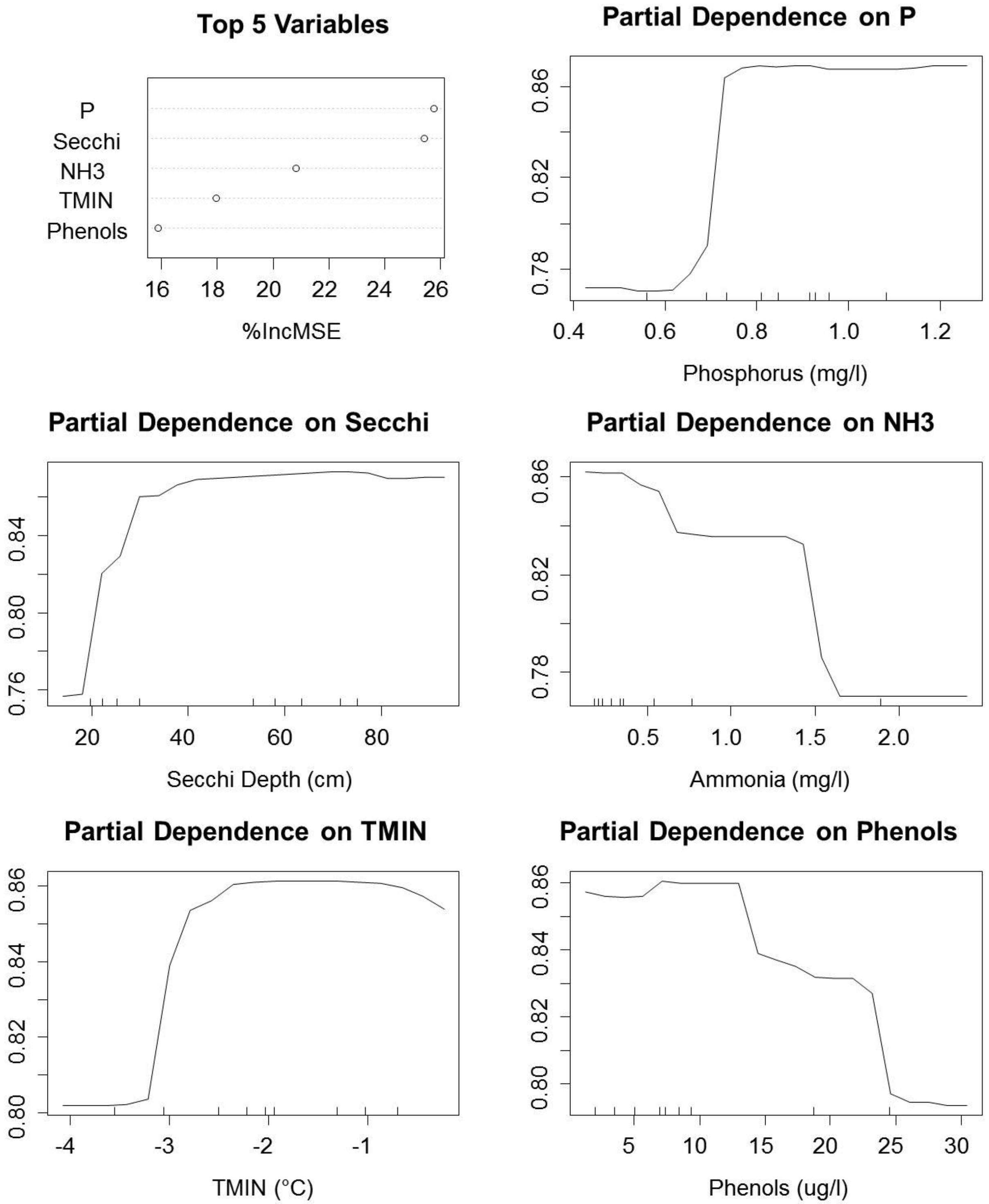
Appendix A: Supporting Random Forests Plots for Chapter 3

Appendix A: Variable importance and partial dependence plots for meaningful random forest models. The y-axis of partial dependence plots represents the response variable, but the range is not meaningful.

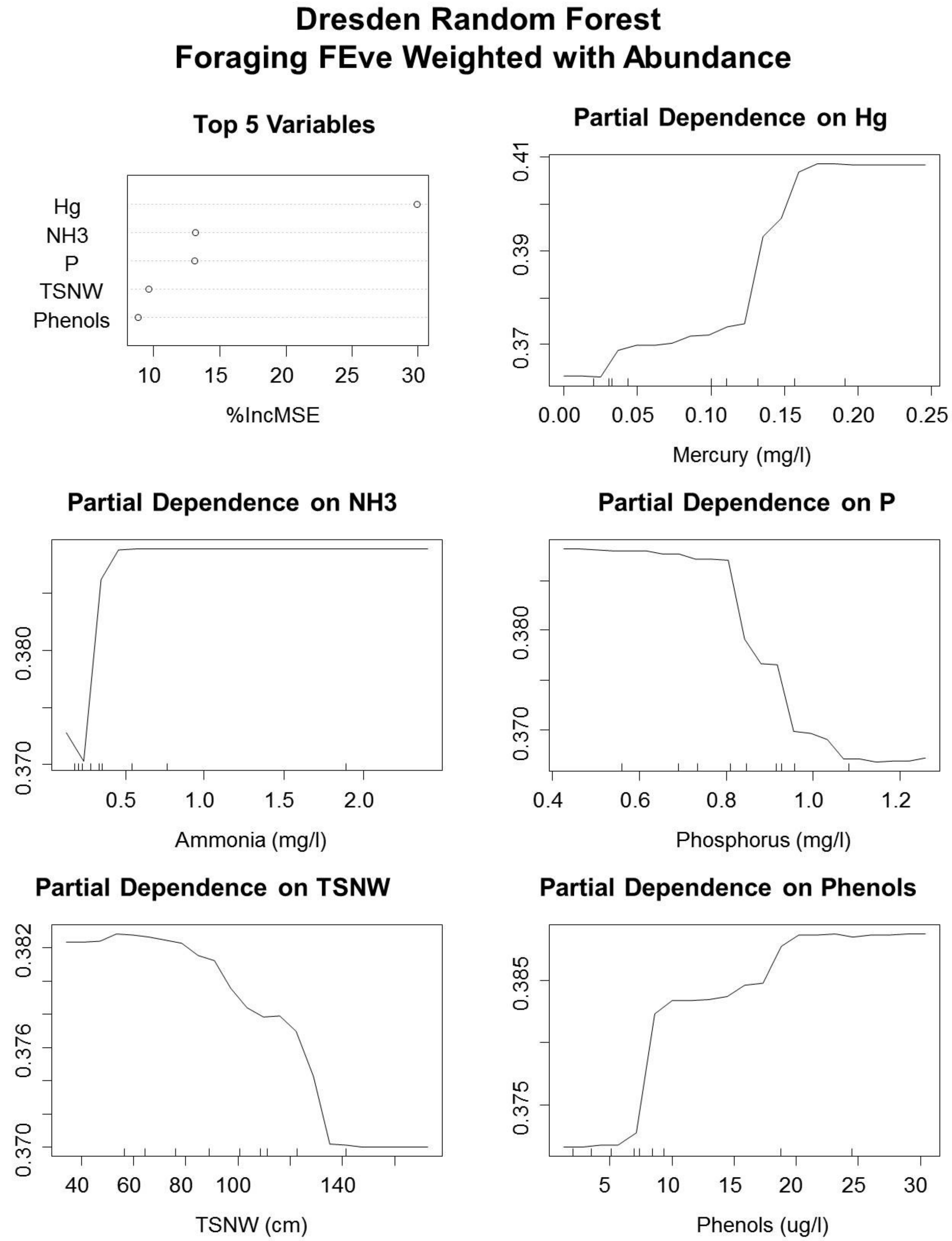


Appendix A: Supporting Random Forests Plots for Chapter 3

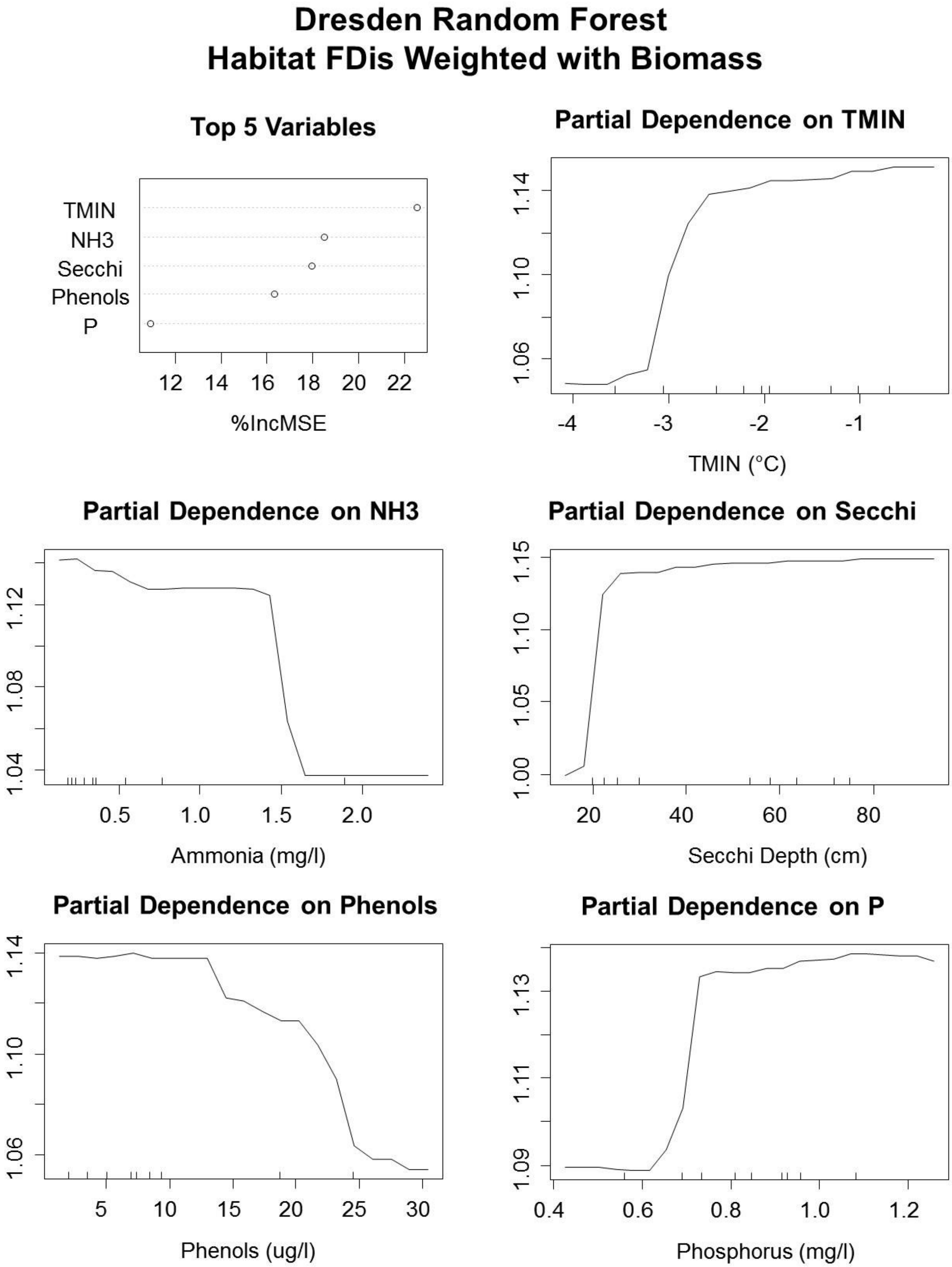
Dresden Random Forest
Foraging FDis Weighted with Biomass



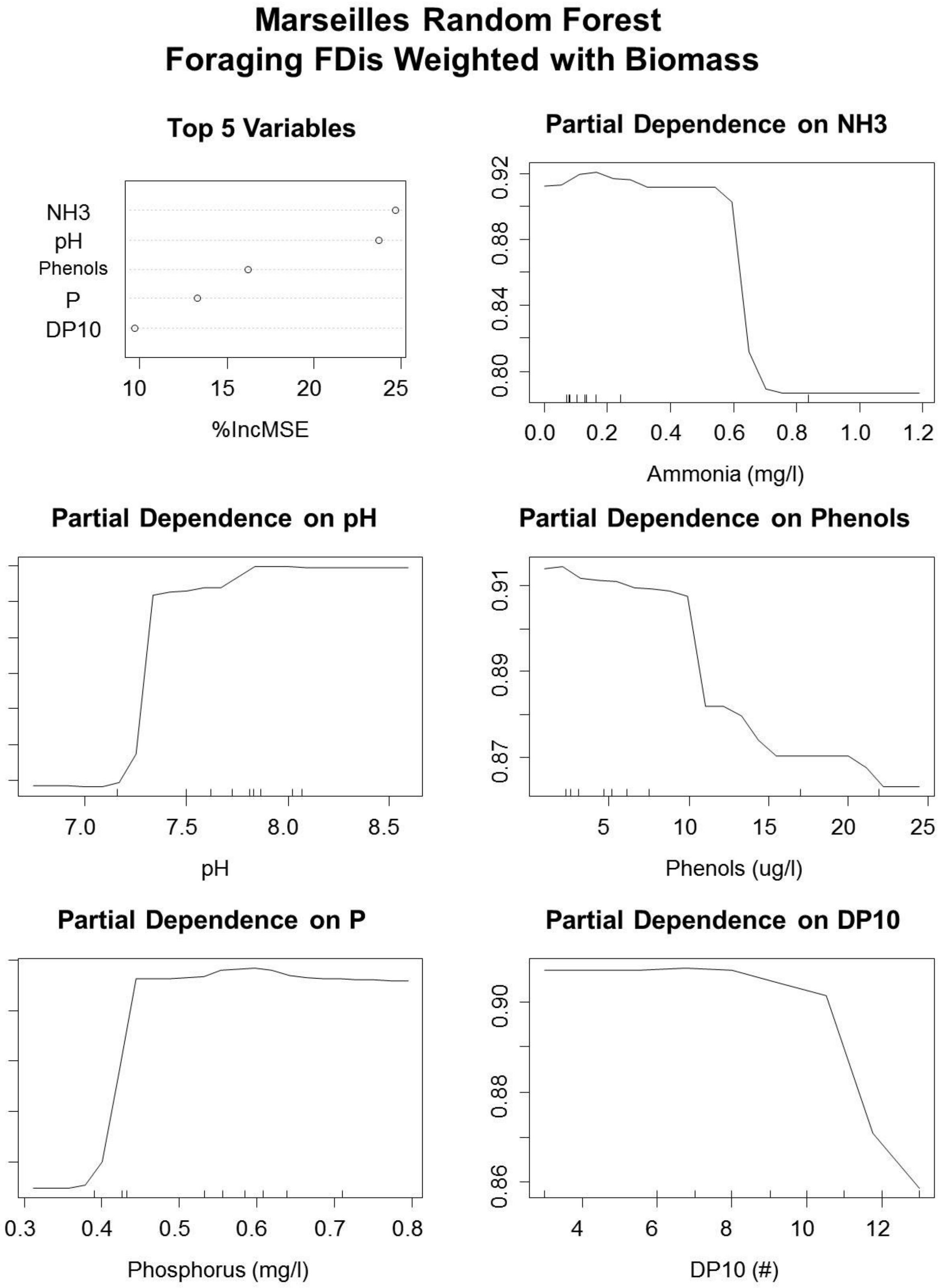
Appendix A: Supporting Random Forests Plots for Chapter 3



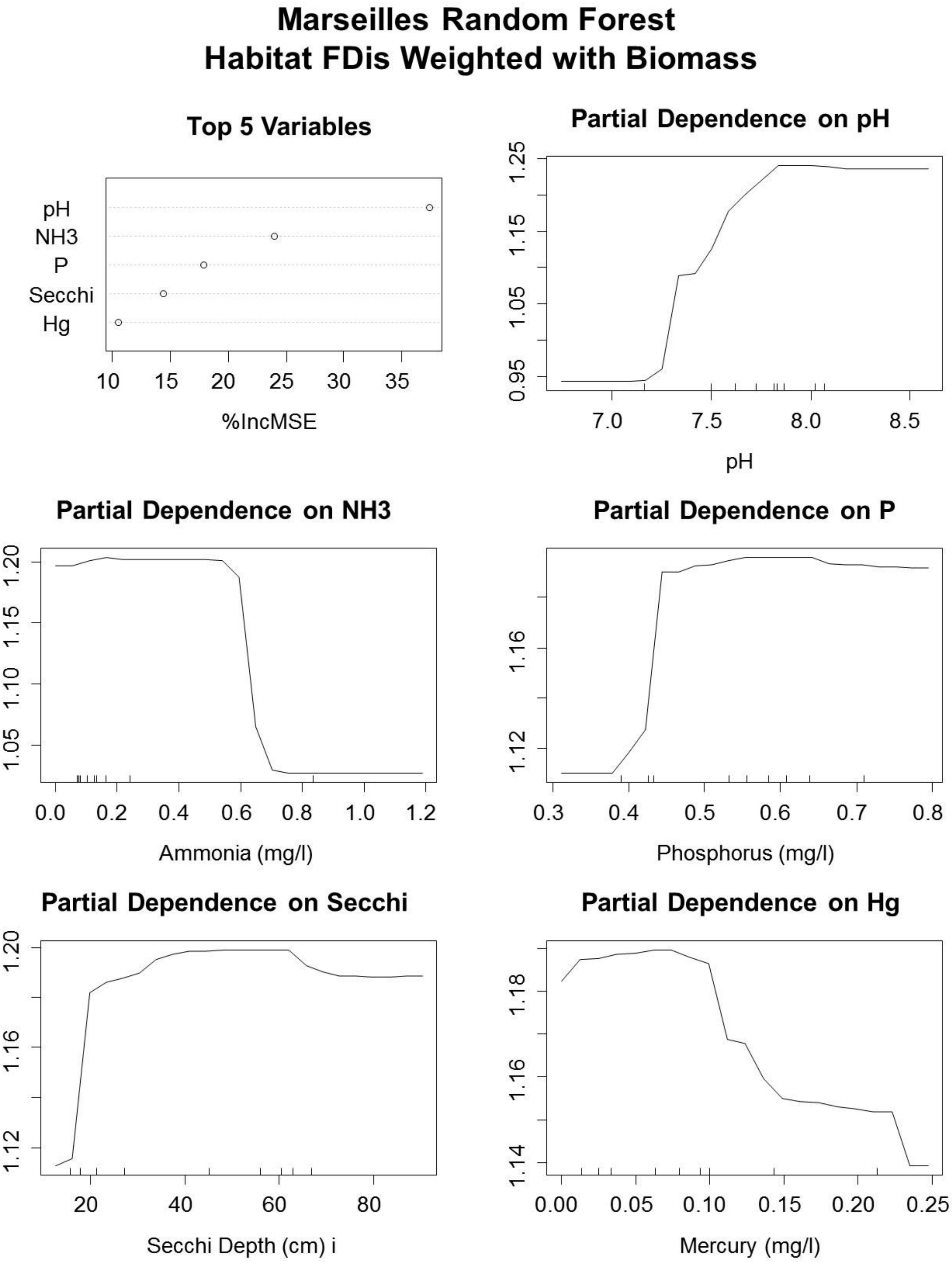
Appendix A: Supporting Random Forests Plots for Chapter 3



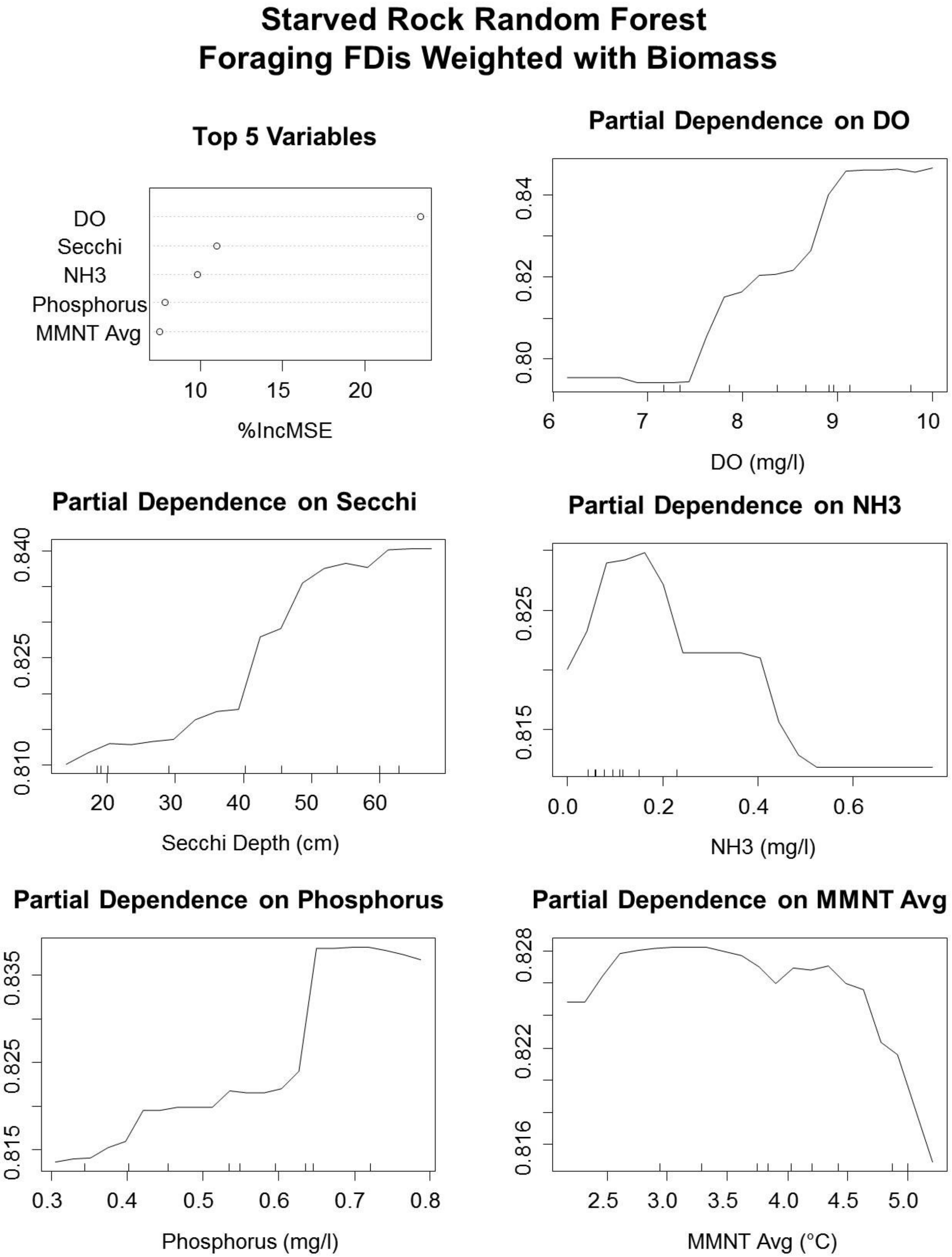
Appendix A: Supporting Random Forests Plots for Chapter 3



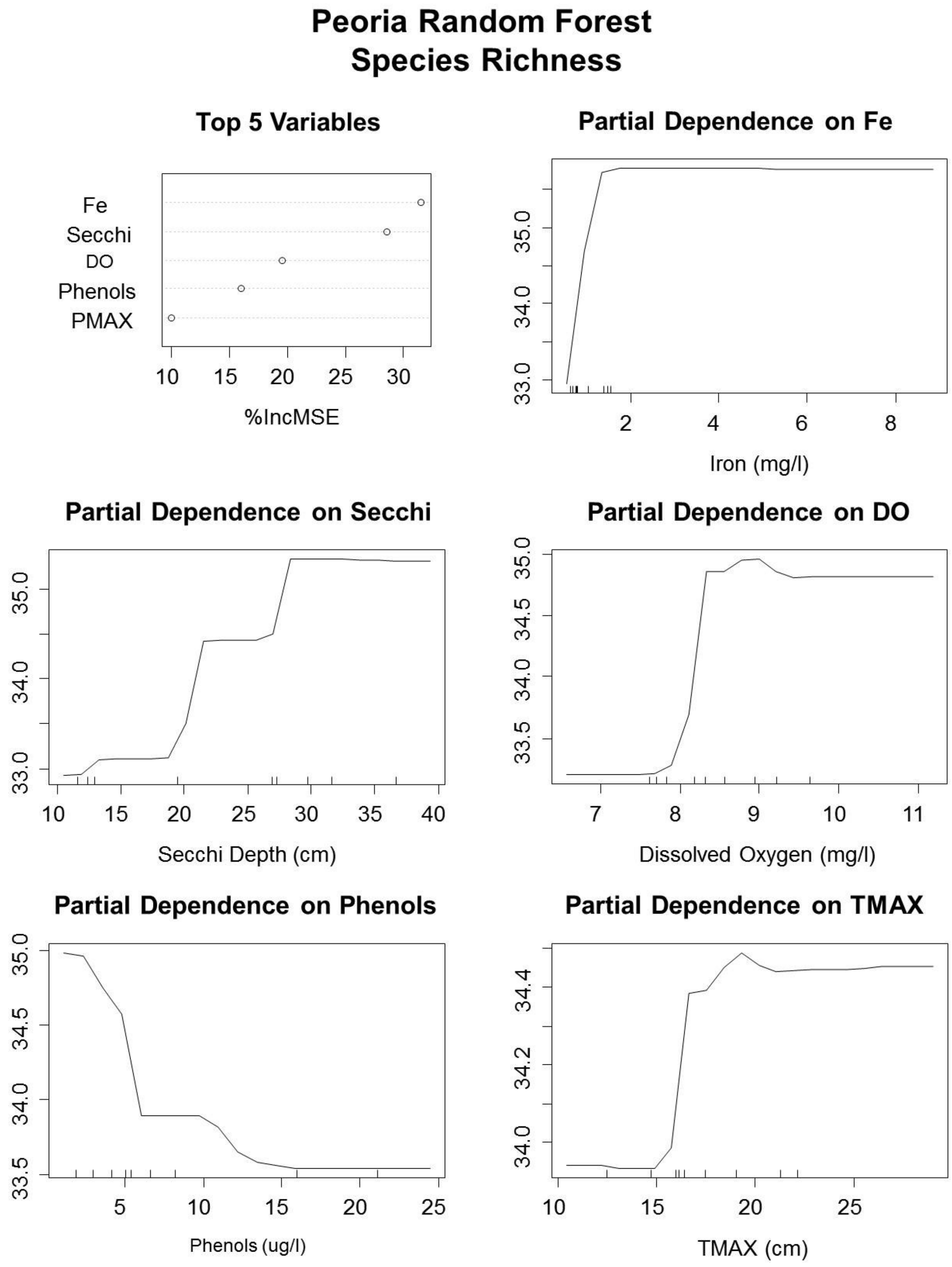
Appendix A: Supporting Random Forests Plots for Chapter 3



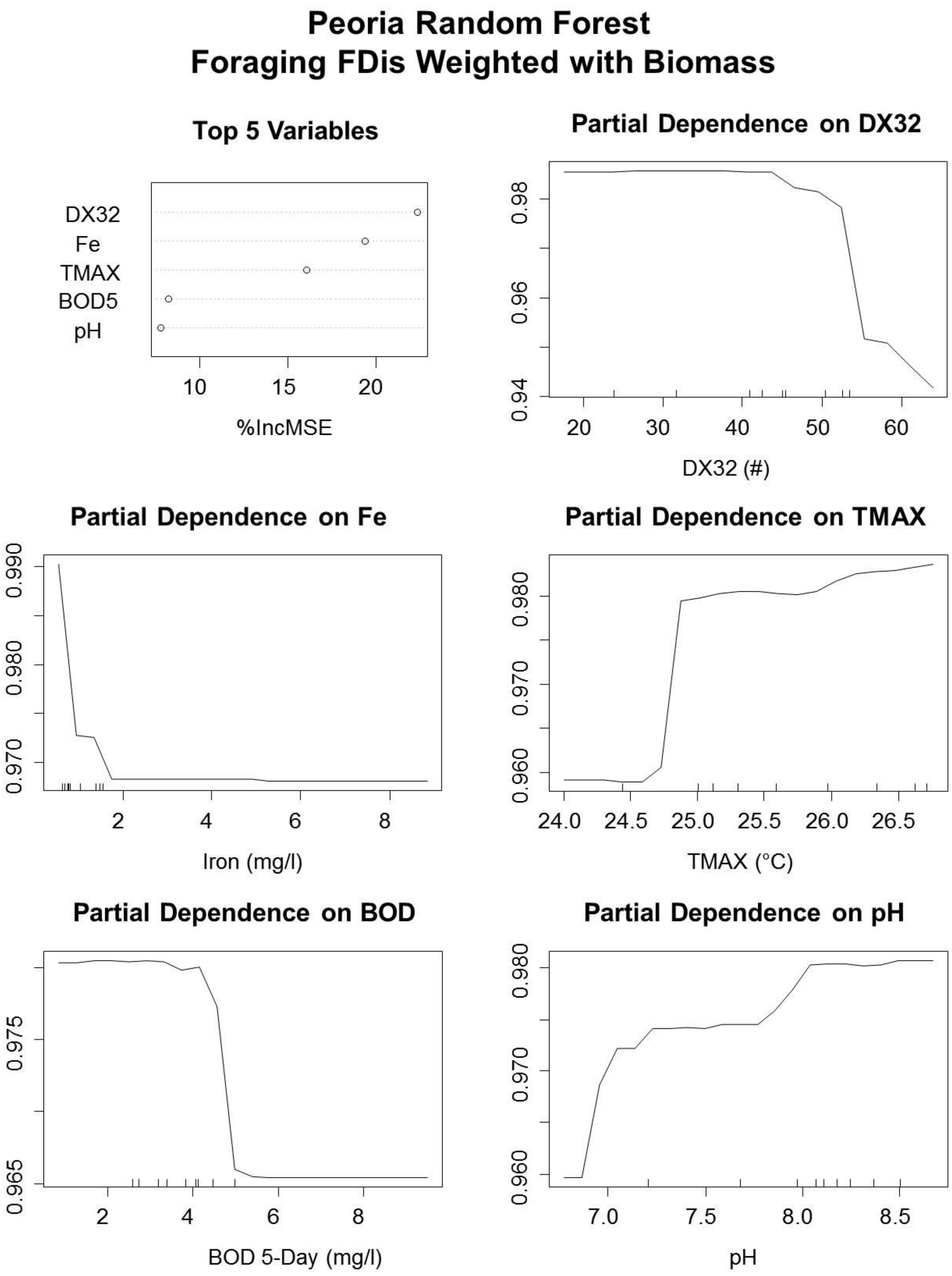
Appendix A: Supporting Random Forests Plots for Chapter 3



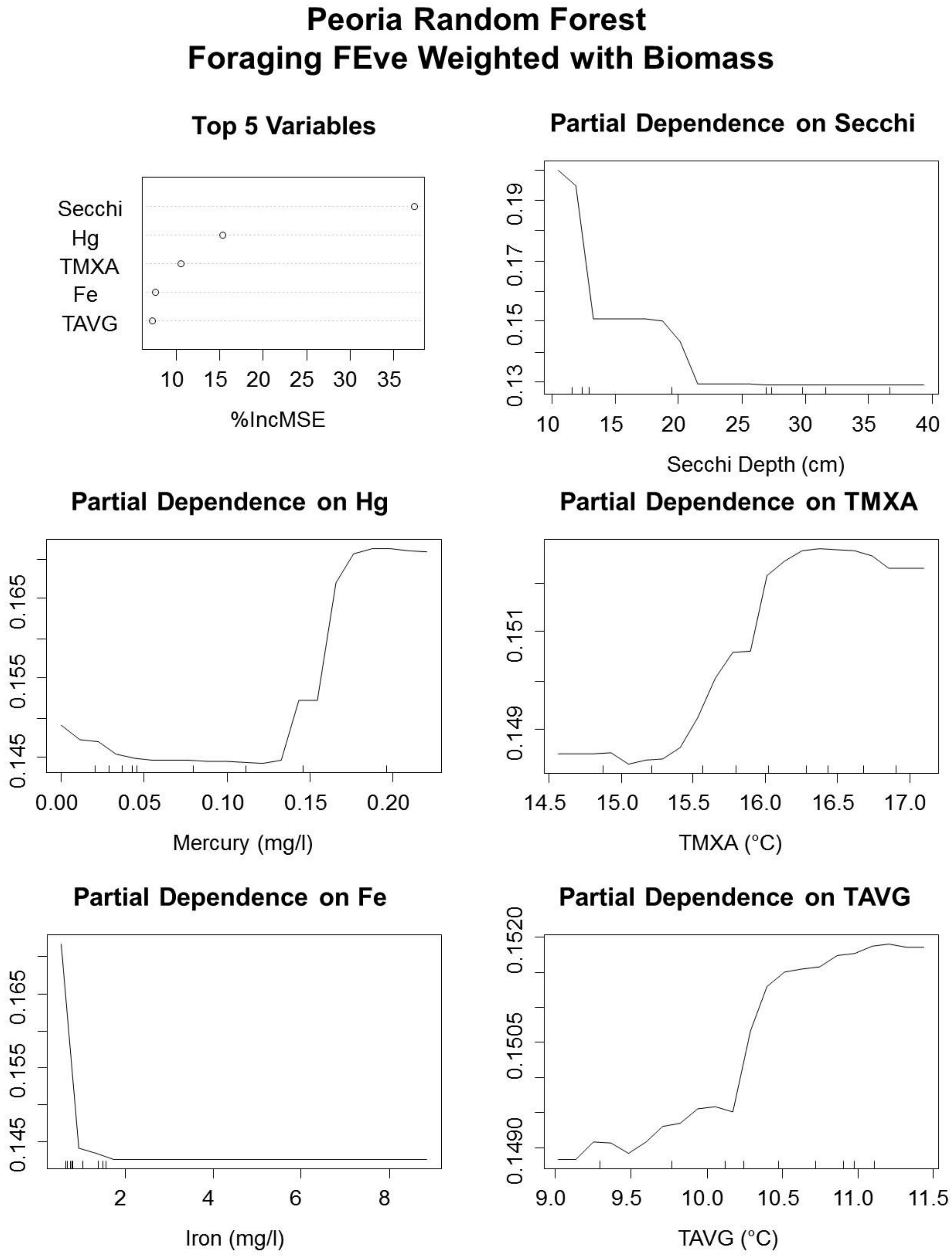
Appendix A: Supporting Random Forests Plots for Chapter 3



Appendix A: Supporting Random Forests Plots for Chapter 3



Appendix A: Supporting Random Forests Plots for Chapter 3



Appendix A: Supporting Random Forests Plots for Chapter 3