# Assessing Water Body Ecological Indicators using Time Series Analysis and Physics-based Modeling Approaches

by

Natalia Bhattacharjee (nèe Shim)

(Under the Direction of Ernest W. Tollner)

Abstract

This dissertation goal is to present a comprehensive time series analysis along with a physics-based model of a stream that can yield expanded ecological data over time which can be further studied with time series analysis. The time series component was primarily developed using available data from a windrow composting operation and runoff collection pond. Descriptive time series analysis (spectral analysis) and detailed recurrent neural network modeling (which requires training and validation data) were performed. Physical stream models and subsequent environmental flow analysis models were developed as aids for water resource management in the Ocmulgee and the Middle Oconee Rivers where we develop a platform useful for evaluating trade-offs between ecological impacts and economic development. We then selected one of the four environmental flow conditions for time series analysis. The methodology and approach set in this work may be adapted to inform environmental flow analyses in other study sites.

Index words: Water Quality, Windrow Composting Pads, Environmental Flows, Habitat Analysis, Hydraulic Modeling, Time Series Analysis, Middle Oconee River, Ocmulgee River

# Assessing Water Body Ecological Indicators using Time Series Analysis and Physics-based Modeling Approaches

by

Natalia Bhattacharjee (nèe Shim)

Integrated BS and MS, Omsk State Technical University, 2006

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2017

Assessing Water Body Ecological Indicators using Time Series Analysis and Physics-based Modeling Approaches

by

Natalia Bhattacharjee (nèe Shim)

Approved:

Major Professor: Ernest W. Tollner

Committee: Jason K. Christian
Nicole Lazar
Abhyuday Mandal
Todd C. Rasmussen

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2017

# Assessing Water Body Ecological Indicators using Time Series Analysis and Physics-based Modeling Approaches

Natalia Bhattacharjee (nèe Shim)

July 24, 2017

# Dedication

This dissertation is dedicated to my family and my husband for providing their unconditional love, support and motivation.

# Acknowledgments

I would like to express my gratitude to my advisor, Dr. Ernest W. Tollner, for giving the opportunity to pursue Ph.D. degree in Engineering at the University of Georgia (UGA). I am very thankful for his guidance and support throughout the years at the graduate school. He has provided invaluable advice throughout my research, encouraged to view the work from an interdisciplinary perspective and supported me in taking courses across various departments at UGA. The past five years of graduate studies has offered the right environment for my professional, academic and personal growth.

I would like to thank my committee members, Dr. Jason K. Christian, Dr. Nicole Lazar, Dr. Abhyuday Mandal and Dr. Todd C. Rasmussen, for providing their suggestions, insightful comments and expertise. I am very thankful for their guidance, patience and time. Additionally, I am grateful to Dr. Kyle S. McKay for assisting in formulating the research framework and providing his expertise in the research area of environmental flows. I also want to thank Dr. Laurie Fowler, Dr. Susan Wilde, Dr. Marguerite Madden for their help and discussions.

Moreover, I am thankful to the UGA-Archway Partnership Program and Carl Vinson Institute of Government for the Public Service and Outreach Graduate Assistantship Award during 2014 – 2015 years. Special thanks to Dr. Michelle Elliot, operation coordinators and Archway professional - Hawkinsville / Pulaski County, for her time answering my questions, providing feedback, helping to improve my communication skills. The discussions helped to put this work into the perspective to community needs. I would like to thank Georgia Water Resource Institute (GWRI) for the opportunity to participate in the State Water Resources Research Institute (WRII) Program supported by U.S. Geological Survey (USGS) and for awarding Research Grant 2015 – 2017 and supporting this research. I am thankful to Dr. Ernest W. Tollner, Dr. Todd C. Rasmussen and Dr. Abhyuday Mandal for their feedback and help in the proposal preparation and submission.

Also, I am thankful to the Center of Geographical Research at UGA for the opportunity to participate in NASA DEVELOP National Program and to gain more knowledge and experience in working with geospatial analysis, earth science data and remote sensing techniques. Additionally, I thank Dr. Nathan and staff in Bioconversion Center for sharing long-term time series data of Windrow Composting Pad operation. Special thanks to Kenneth W. Swinson and Joshua R. Willis for assisting in bathymetry data collection. I would also like to thank Joshua Willis, Justin Wale and Kallie Craft in taking part in this research. Thanks to Ike Astuti, Manasi Parkhi, Dorris Scott, Wen Gu, Yaping Yang, Mayuri Perera, Eliana Mozo Reyes, Megan Filkowski, Tiffany Vidal, Caren Remillard, Christopher Cameron, Emad Ahmed, Kamala Kanta Sahoo, Joshua Willis, Ken Swinson and many others for their help in this journey. I am thankful to Hansol Lee, who became one of my best friends during the graduate studies at UGA, for her amazing support, encouragement and friendship.

I am so grateful to my parents who have been sharing my experience in this journey from a long distance and supporting me in studying and working abroad for more than ten years. I am very thankful for all their patience and love. Special thanks to my cousin Elena for supporting and sharing my experience in pursuing scholarship and research opportunities abroad.

And last, but not least, to my husband and life partner, Tapomayukh Bhattacharjee, for his encouragement to initiate this journey and for believing in me. I cannot thank him enough for his continuous support, love and care; and for being there when I needed most.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Background and Justification

"We won't have a society if we destroy the environment", as Margaret Mead once commented, profoundly indicates how environmental concern is embedded in public life. If we want to help the society, we should solve environmental problems and develop technologies for sustainable management of natural resources. Sustainable development is one of the main goals of environmental management (Dale, 2003a). Environmental managers need to ensure that they make decisions without compromising the ability of future generation to meet their needs (Comum, 1987). The decision-making process is usually supported by ecological modeling that uses evolving technologies and approaches. However, the increasing environmental pressures require refinement of management strategies and use of an integrated modeling framework (Dale, 2003b).

This dissertation develops integrated solutions for hydrological, water quality and ecological applications using hydraulic modeling as well as spatial and time series analysis. The work focuses on building a foundation of developing new knowledge for the improvement of water resource management.

Quantitative and qualitative hydrologic forecasting is crucial in the field of water resource and environmental engineering. Information from the forecasts can be used in order to control a system, anticipate and prevent future disasters. However, accurate quantitative and qualitative hydrologic forecasting is a challenging task due to uncertainty involved in ecological modeling, uncertainty in hydrology, as well as due to climate change and variability. Traditionally, approaches of understanding and describing system behavior are divided into two groups: physics-based and data-driven modeling. Physical modeling starts from theory wherein data-driven modeling develops from data. Both physical and data-driven types of model are widely used for hydrological predictions and both have advantages and disadvantages. Most models that have been developed require much site specific data collection and, once commercialized or publicized, provide a limited access to the process model themselves. This research focuses on integration of both physical and data-driven approaches for purposes of hydrological forecasting in the context of quantitative as well as qualitative predictions. This topic has been studied by researchers (Shafroth *et al.*, 2010), (Jones-Farrand *et al.*, 2011), (Jones, 2014), (Poff *et al.*, 1997) and (Richter *et al.*, 2006) but there are still many questions that need to be answered.

This dissertation includes a case study of the windrow composting pad management that enables operators to anticipate conditions when water quality concentration exceeds regulatory thresholds. The detailed sensitivity analysis of a recurrent neural network can allow a better understanding of water quality dynamics of collected runoff and assist in identifying strategies for better management of windrow composting systems. This time series modeling approach can be applied (although not done in this work) on larger problems such as river flow management to help anticipate adverse environmental flow conditions and water quality issues.

Another two case studies are related to water resource management in the Ocmulgee and the Middle Oconee Rivers where we examine trade-offs in water management between ecological impacts and economic development to meet future demands in Georgia. Population in the Middle Ocmulgee Region of Georgia is projected to double by 2050 (increasing from $567,728$ in 2010 to $1,180,000$ by 2050) which will involve water withdrawals increase by 38% and wastewater return flows increase by 62% by 2050 (RWP, 2011). This will require better water management and supply system in the future (RWP, 2011). A comprehensive analysis of environmental flow schemes represents a key step in ensuring adequate water availability to meet increasing human needs while minimizing adverse impacts on aquatic ecosystems. The methodology and approach set forth in this work may be easily adapted to inform environmental flow analyses in other study sites.

## 1.2 Statement of the Problem

### Composting Pad Management

Continuous industrialization and increasing population are resulting in many water quality problems. Much organic waste originates from municipalities in the form of yard waste and municipal biosolids. Windrow composting systems are alternatives for degrading organic pollutants while preserving opportunities for beneficial reuse. Extension/outreach personnel in Georgia estimate that there are some 40 to 50 windrow composting facilities in Georgia. Having a composting option is important to certain industrial sectors of interest to rural Georgia. Windrow composting sites are usually open field areas exposed to rainfall. Many are not situated on a permitted sewage treatment location and thus require liquid waste permitting.

The surface of a composting pad is typically made up of crusher-run rock that provides a sturdy and pervious base for the compost. The rock media overlays a compacted clay. The surface runoff and infiltrated percolant components are typically directed to a holding pond at the downslope end of the pad. Pad runoff is highly regulated due to pollution potential (Wilson *et al.*, 2004, Kalaba & Wilson, 2005). Therefore, the amount of runoff and water quality must be considered when designing a pond and collecting runoff in order to prevent the discharge of organic pollutants. One approved approach for discharging the pond effluent is through an approved land application system (LAS). Factors affecting water quality in the pond are of key interest and need to be better understood in order to better manage LAS pond. In late summer months, water quality of pond effluent often exceeds regulatory EPA limits for reasons that need further study.

### Sustainable Development

With increasing population in Middle Georgia, water demand and wastewater generation are dramatically increasing. This requires refining water management strategies in order to meet future demands and support economic development of the region. River flow in both river basins is affected by Wallace Dam and Lloyd Shoals Dam that are located near Oconee Lake and Jackson Lake, respectively. Therefore, our study compares the two river basins and examines the ecological effects of alternative water management practices in both the Oconee and Ocmulgee river basins (Figure 1.1). These rivers harbor high aquatic biodiversity, and protecting these species is of high priority. Thus, there is a need to examine trade-offs in water management between ecological impacts and economic development and to analyse how reservoir operation can influence local hydrology and fish habitat. It is important to support development of environmental flow regulations in the Oconee and Ocmulgee river basins and contribute to the improvement of local water management and planning to ensure sustainable development and to meet future generation water demands. Additionally, this work supports National Park Initiative (Figure 1.2). On $30^{th}$ January 2017, the bill for expansion of the Ocmulgee National Monument into Georgia's first national historic park, passed U.S. House of Representatives (http://www.macon.com/news/local/article129705059.html). So far, there is no National Park located in Georgia (Figure 1.3).

## 1.3 Study Sites

We examine two river basins: Oconee River and Ocmulgee River (Figure 1.1). These rivers are adjacent basins and major tributaries that join to form the Altamaha River. Both river basins offer aquatic diversity and protecting these species is of high priority for the region. Thus, it is essential to study flow regimes and trade-offs involved in water management to take into account the impact that the development will have on the environment. The research is focused on development of a framework for reaching a balance between reservoir power generation and environmental flow that are beneficial for recreational activities, such as fishery, canoeing/kayaking. We analyze how reservoir operation can influence local hydrology and fish communities to gain necessary information for a sustainable ecosystem along Oconee and Ocmulgee River.

We use Hydrologic Engineering Centers River Analysis System (HEC-RAS) software developed by US Army Corps of Engineers to model hydraulics of the river. We also use HEC-GeoRAS, another software developed by US Army Corps of Engineers. HEC-GeoRAS allows us to import data from Geographical Information System (GIS) software to HEC-RAS. Our analysis of how reservoir operation can influence local hydrology and fish communities can provide information for sustainable ecosystems in Oconee and Ocmulgee rivers.

The study area is a small reach of the Middle Oconee river near Ben Burton park in Athens, Clarke county in Georgia. The length and width of Oconee river region are estimated as 1200 ft and 150 – 200 ft, respectively. Hourly time series discharge data were obtained via website of the U.S. Geological Survey (USGS) water resources which operates streamflow monitoring station near the study area (Gage number 02217500). The mean daily flow discharge based on 81 years of records is about 2 L/s (286 cfs). Another study area is a reach of Ocmulgee river near Macon, Bibb county of Georgia state. The length and width of Ocmulgee river are 5000 ft and 500 – 800 ft. The USGS gage number close to this study area is 02213000. The mean daily flow discharged based on 88 years of records is around 40 L/s (1440 cfs).

## 1.4 Objectives

Figure 1.4 illustrates the main idea of our work where we use physics-based and data-driven approaches to develop integrated ecological models in order to improve water resource management. These studies incorporate temporal data of weather, water flow, water quality and spatial data (bathymetry data, LiDAR data, ASTER DEM data) to link hydrological, hydraulic and ecological models and support decision-making process involved in environmental management. Please note that work related to calibration and validation is not included in this document as it is a part of my MS thesis towards MS degree in Statistics.

The following research questions are addressed in this dissertation:

(a) How trends in time series indicate likelihood of approaching exceedance of selected measures?

(b) What factors affect dynamics of water quantity and water quality time series?

(c) What factors lead to better management of hydrologic and water quality parameters?

(d) How we may use sensitivity and uncertainty analysis to streamline the process for building integrated deterministic and stochastic models?

(e) How far downstream the flow regime and fish recruitment are affected along the Ocmulgee reach from Macon to Hawkinsville?

(f) How we may estimate economic and ecological trade-offs under different flow regime scenarios?

(g) How we can evaluate "environmental flows" for a small river?

This dissertation consists of literature review chapter (Chapter 2), three chapters (manuscripts) reporting individual case studies and a summary/conclusion chapter. The objectives given above were addressed by looking at three case studies and each study is described in a separate chapter (Chapter 3-5). Chapter 3 focuses on water quality prediction of runoff from UGA windrow composting pad. Chapter 4 investigates alternatives on water management and includes habitat analysis on a small reach of the Middle Oconee river near Ben Burton Park. Chapter 5 extends the previous work by examining how well this approach can be generalized and applied to other sites and focusing on Middle Ocmulgee region area. We conclude with time series analysis of selected ecological indicators. The common goal of the three studies is to provide insights for a better management of ecological system and sustainable development of a region using as much available data as possible. Chapter 6 gives summary, conclusion and future research.


Figure 1.1: Research Study Sites: Oconee and Ocmulgee Rivers.

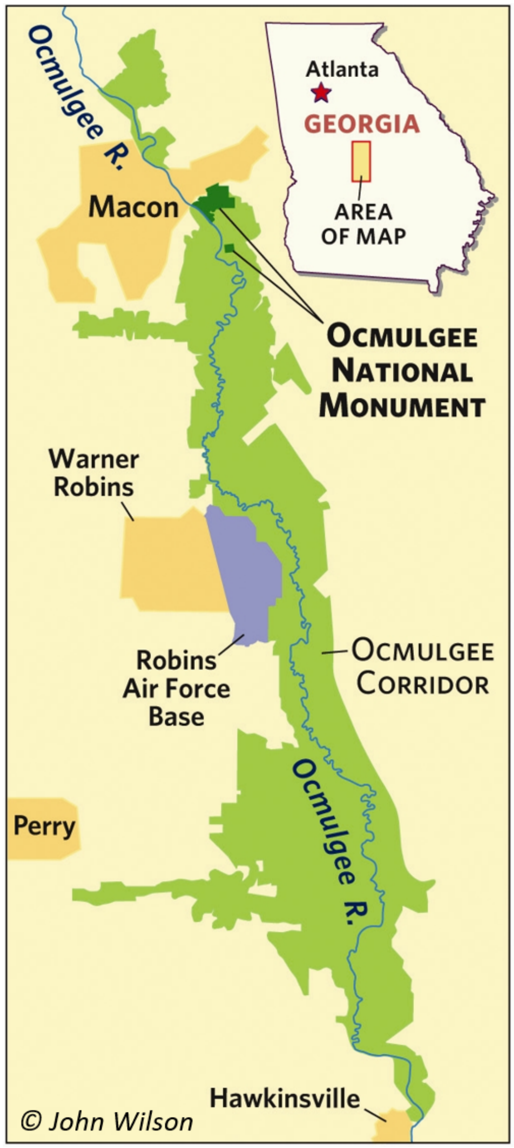



Figure 1.2: The Ocmulgee National Park and Preserve Initiative.


Figure 1.3: National Parks in the United States.

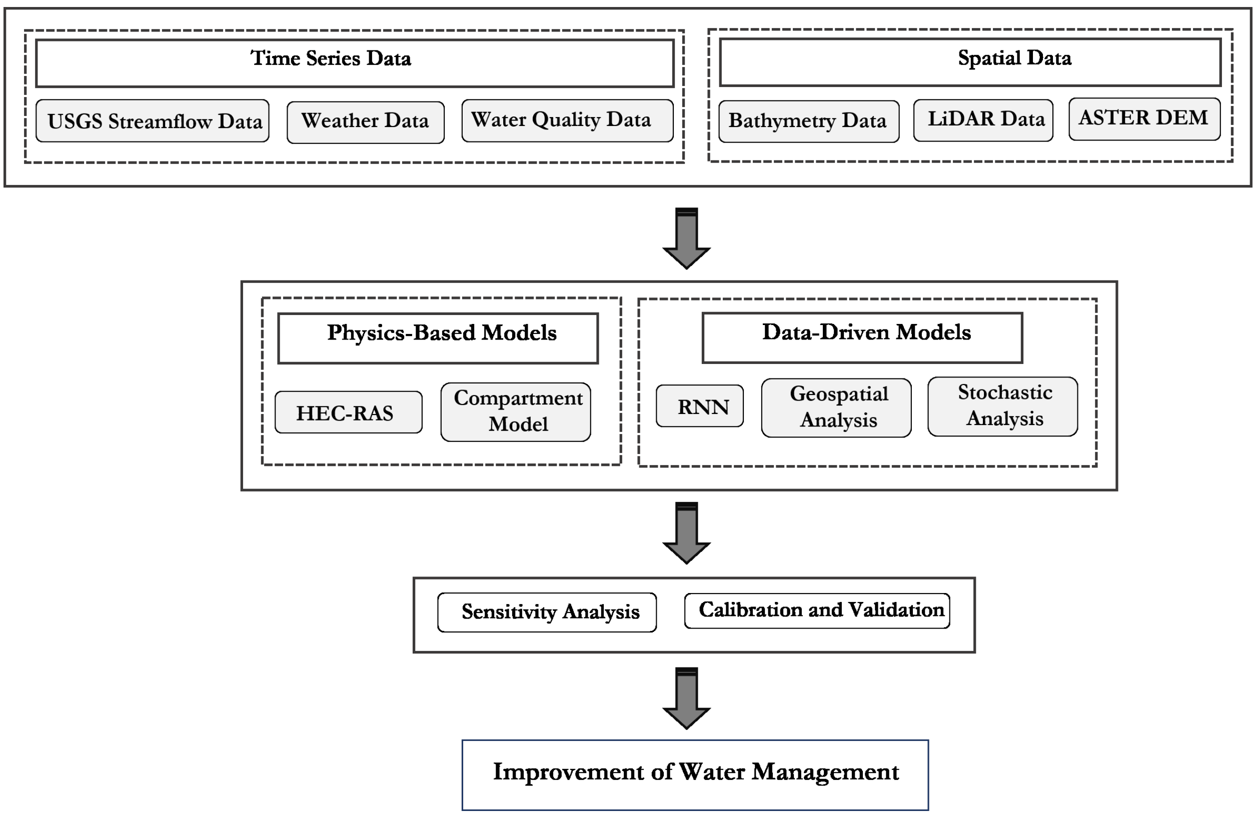



Figure 1.4: Conceptual visualization of the research framework.

# Chapter 2

# Literature Review

The literature review is organized as following: Section 2.1 covers related work on windrow composting, its management and water quality modeling; Section 2.2 describes water management, environmental flows and hydraulic modeling; Section 2.3 includes previous studies on habitat suitability and ecohydrology.

## 2.1 Water Quality Prediction

Much organic waste originates from municipalities in the form of yard waste and municipal biosolids. Windrow composting systems are alternatives for degrading organic pollutants while preserving opportunities for beneficial reuse. However, windrow composting sites are usually open field areas exposed to rainfall (Dorahy *et al.*, 2009). Typically, windrow composting systems are located in remote areas and runoff cannot be directly treated in a wastewater treatment plant. Therefore, the runoff must be collected into a pond prior to its release from the facility (Kalaba *et al.*, 2007). The runoff from these facilities does not meet water quality standards and is highly regulated due to pollution potential. Therefore, the amount of runoff and water quality must be considered when designing a pond and collecting runoff in order to prevent the discharge of organic pollutants.

One approved approach for discharging the pond effluent is through an approved land application system (LAS). The main advantage of compost pad application is avoiding adverse environmental effects associated with uncontrolled organic pollutants, such as leachate and methane production (Dorahy *et al.*, 2009). State regulations prefer municipal waste treatment of this waste stream and provide little direct guidance to windrow compost operation not located on a municipal treatment plant site. There remain some significant science issues in the design of water catchments in land application systems. In this study, we will focus on factors affecting water quality of stored runoff in order to improve understanding of pollutant dynamics in compost runoff and catchment.

Many researchers tried to estimate and predict the runoff by suggesting different computer simulation models such as Soil-Plant-Atmosphere-Water (SPAW), Groundwater Loading Effects of Agricultural Management Systems (GLEAMS), Soil Water Assessment Tool (SWAT) that rely on a constant curve number. However, issues in runoff predictions using the constant curve number method occurred due to various compost materials and surface types ranging from concrete to gravel. Tollner and Das (2004) suggested a dynamic curve number method for a better estimation of runoff from windrow compost pads when compared to a constant curve number method. Kalaba et al. (2007) provided evidence that water transmission delay from the compost pad can lead to discrepancies in runoff estimation. Duncan et al (2013a) revisited the curve number method in an analysis including data since 2004 and found excellent event-based results. They went on to develop the dynamic simulation because flow rate modeling was possible. Therefore, a thorough study of a number of factors affecting the runoff from a compost pad is required to predict the amount of runoff from the rainfall events. In a recent study, Duncan et al. (2013b) tried to overcome the difficulties and extended knowledge of rainfall-runoff relations for the windrow composting pad. Duncan et al. (2013b) proposed a compartmental dynamic model for estimating the amount of runoff and developed an instantaneous unit hydrograph (DeCoursey, 1966, **?**, Huggins & Burney, 1982) based on the model.

Application of neural networks showed extreme growth in different research areas, including water management. Nourani and Fard used artificial neural networks in order to model evaporation process (Nourani & Fard, 2012). Garbrecht used ANN for rainfall-runoff modeling (Garbrecht, 2006). Toth et al. applied neural networks for flood forecasting (Toth *et al.*, 2000). In the field of water quality modeling, Schmid and Koskiaho applied static ANN known as a multilayer perceptron (MLP) type of neural network for the modeling of dissolved oxygen in a wetland pond (Schmid & Koskiaho, 2006). Milot et al. studied contribution of MLP (or feedforward neural network (FFNP)) for modeling trihalomethanes occurrence in drinking water (Milot *et al.*, 2002). Verma et al. implemented MLP network for predicting total suspended solids (TSS) in wastewater (Verma *et al.*, 2013). Suen and Eheart evaluated performance of radial basis function (RBF) neural network (a type of statistical neural network) for modeling nitrate concentrations in a river (Suen & Eheart, 2003). Singh et al. constructed FFNP for computing dissolved oxygen (DO) and biological oxygen demand (BOD) of the river water (Singh *et al.*, 2009). Also, they adopted FFNP approach for predicting BOD and chemical oxygen demand (COD) levels in wastewater treatment plant effluent (Singh *et al.*, 2010).

## 2.2 Water Management

$7Q10$ is the annual minimum 7-day average flow that occurs on average once every 10 years (Fisher & Thompson, 2003). Thus, there is 10% probability that there will be a lower flow in any given year. $7Q10$ is reported to be 1.1 L/s (37 cfs) for the Middle Oconee river near Athens (gage number 02217500) (McKay, 2014). $7Q10$ is reported to be 8.6 L/s (305 cfs) for the Ocmulgee river near Macon (gage number 02213000) (Gotvald, 2016). From 2001, Department of Natural Resources (DNR) adopted Monthly $7Q10$ regulation as a monthly minimum flow policy. Monthly $7Q10$ is a statistical figure that reflects the lowest 7-day average flow for each calendar month with a 10 year recurrence interval (www.gaepd.org).

With increasing human population and industrial development in Middle Georgia, future increases in water demand and wastewater generation will further strain riverine ecosystems in the region (RWP, 2011). This requires refining water management strategies to meet future demands and support economic development of the region. The Middle Georgia region relies equally on both surface water and groundwater sources. It contains three river basins: Flint, Oconee and Ocmulgee. In this work, we are interested in water resources management of the Middle Oconee and Ocmulgee Rivers. Over the years, river course may change based on land changes, urbanization and flood events (Lane & Richards, 1997). However, the river characteristics (bathymetry) may not change much; therefore, in this study, we assumed the bathymetry of the stream to be unchanging. One could estimate sediment transport using HEC-RAS to look at the erosion and deposition processes. However, this was not considered in these case studies.

A wide variety of strategies exists for identifying environmental flow targets and thresholds, which have been reviewed elsewhere e.g. (Jowett, 1997, Tharme, 2003, McKay, 2013). While more complex methods exist, simple hydrologic thresholds remain extremely common in practice. Hydrologic indices calculated from historically observed discharge data at daily, monthly, or annual time-scales typically form the basis for these operational rules. Minimum flow levels set a river discharge (or stage) below which water may not be withdrawn, and these techniques are extremely common in setting regulatory thresholds at annual or monthly timescales. Percent of flow methods represent a second simple environmental flow technique, in which a percent deviation from an upstream discharge rate guides withdrawal amounts or reservoir operations. Percent of flow methods, also known as sustainability boundaries (Richter, 2010, Richter *et al.*, 2012), are gaining popularity, in part due to straightforward operational goals and the capacity to preserve natural variability in flow regimes.

Rivers are extremely dynamic systems, and flow regimes often exhibit many sources of variability both within a year (e.g., seasonal periodicity) and between years (e.g., wet and dry years) (Sabo & Post, 2008). Accordingly, a growing body of researchers have pressed river managers to not only manage variability, but manage for variability (Arthington *et al.*, 2006, Poff, 2009, McKay *et al.*, 2016). However, environmental flow thresholds are often set based on "typical" river flow levels (i.e., long-term averages or central tendencies). Historically, river engineering has struggled to cope with the challenges of what discharges to use in design and management, and a common method for incorporating the magnitude and frequency of events is "effectiveness analysis" (Wolman & Miller, 1960). Recently, this family of techniques has been adapted for ecological applications in streams (Doyle *et al.*, 2005, Wheatcroft *et al.*, 2010, McKay *et al.*, 2016).

## 2.3 Habitat Suitability

It is important to quantify several trade-offs that would help to further develop recreational uses along the Ocmulgee River. Therefore, we are interested in quantifying an influence of hydrological historical and future events on hydropower operation that, in turn, affects recreational use. For example, trade-offs between maximum power generation versus flood risk, daily water releases versus recreational minimum and maximum levels (that are beneficial for boating and fishing), reservoir coordination strategies versus environmental requirements, and water allocation versus reservoir minimum levels, can all be evaluated. Previous research shows a relationship between flow and fish biomass based on study for Austrian rivers (Moog, 1993). Figure 2.1 suggests that high ratio between maximum and minimum discharge negatively impacts diversity of fish. Based on U.S. Geological Survey (USGS) data on Ocmulgee river flow near Jackson, GA, the ratio between maximum and minimum discharge is 18.6 which indicates that it puts high stress on fish biomass (Figure 2.1). Researchers also suggest that extended periods of low flow lead to changes in aquatic habitat (Smakhtin, 2001) and investigated whether key-habitat availability can be used as a criterion for indicating potential effects of flow regimes on fish species.

For example, stream fish assemblage data and hydrologic data (34 sites in Wisconsin and Minnesota) were used to test the hypothesis that the organization of fish communities are related to hydrologic variability (Poff & Allan, 1995). That study accounted for environmental factors such as depth, velocity, food availability, and thermal regime and found a strong relationship between hydrologic variability and fish assemblages.

In another study, habitat suitability criteria for nine fishes (shiner and darter species) were developed for Piedmont and Coastal Plain streams in Alabama by Freeman et al. (1997). Criteria analyzed included depth, velocity, substrata type, and cover. This study tested the hypothesis that a higher sample of fish will be seen in optimal habitat as opposed to suboptimal habitat and it tested the transferability of habitat suitability criteria between some streams (Freeman *et al.*, 1997).

Gregory et al. (2002) assembled records of fish species from various research reports (from museums, agencies, databases and field data). Known records and locations were recorded via GIS database and used in order to generate potential distribution maps for fish species. Researchers suggest that local habitats, landscape and water management affect fish assemblages in the Willamette River. Lowlands region of the Willamette River Basin (WRB) in Oregon state, has higher number of fish species when compared to Upland regions (native fish species along with introduced species) because introduced species tend to occupy warmer, low gradient streams (Gregory *et al.*, 2002).

Scientists also looked at relationship between key-habitat availability and fish abundance in the Tallapoosa River, southeastern U.S. river located in Alabama (Bowen *et al.*, 1998). They measured depth and velocity for cross-sectional profiles and using the field data, researchers simulate hydraulic conditions for a range of flows in Physical Habitat Simulation (PHABSIM) model.

For eastern warmwater U.S. rivers, it is not practical to collect habitat data due to high number of present fish species and common habitat patterns. Therefore, more generalized criteria than species-specific habitat suitability criteria should be use for warmwater fishes (Bain, 1995). One of the more general approaches is key-habitat criteria. Bain (1995) introduced five key-habitats with respect to hydraulic terms (shallow-fast (super-critical), deep-fast (run), shallow-slow (riffle), slow-cover and shallow-coarse) given in Table 2.1.

Additionally, researchers studied alligator habitat response to environmental factors, such as temperature. Many researchers consider temperature as a critical factor relevant to growth of alligators (case studies in Louisiana, North Carolina and Florida Keys) (Lance, 2003, Joanen & McNease, 1989, Seebacher *et al.*, 2003). Temperature below 16°C is described as a condition when alligators stop eating and stop growing (Joanen & McNease, 1989, Seebacher *et al.*, 2003). Joanen and McNease, 1989 refer the months when the temperature above 16°C as growing month.

Table 2.1: Habitat suitability criteria (Bain, 1995).

| Key Habitat | River Depth | Flow Velocity |
|---|---|---|
| Shallow-Fast | ≤ 35 cm | ≥ 55 cm/s |
| Deep-Fast | ≥ 35 cm | > 45 cm/s |
| Shallow-Slow | > 35 cm | < 35 cm/s |
| Shallow-Coarse | < 35 cm | Gravel or larger substrata present |
| Slow-Cover | structure present | ≤ 20 cm/s |

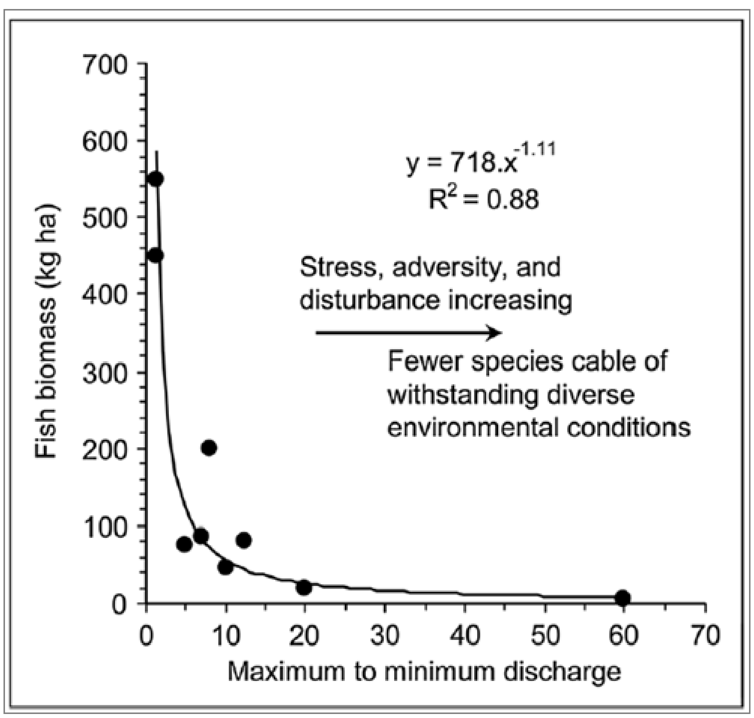



Figure 2.1: Association between hydropeaking flows and fish bisomass (Moog, 1993)

# Chapter 3

# Improving management of windrow composting systems by modeling runoff water quality dynamics using recurrent neural network[1]

[1]Bhattacharjee N. V. and Tollner E. W. 2016, Ecological Modeling 339: 68-76. Reprinted here with permission of publisher.

# Abstract

The recurrent neural network is a tool that can provide valuable insights when forecasting future likelihood of events using dynamic time series. One of the challenging research problems is to extend the black-box modeling into white-box modeling in order to gain insights into the physics-based processes. Sensitivity analysis has shown a great contribution in overcoming this challenge. The main objective of this study was to perform a detailed sensitivity analysis of recurrent neural network in order to identify parameters that are important for predicting water quality constituents. We used a windrow composting pad located at UGA Bioconversion center, Athens, GA, USA as our study site. Runoff from windrow composting pad was collected in order to prevent the discharge of organic pollutants. We used time series data from nine years of precipitation, temperature, pond volume, material volume on the pad, total suspended solids (TSS), biological oxygen demand (BOD) and nitrate ($NO_3$) concentration levels of stored runoff in the collection pond. Previously, we applied recurrent neural network for predicting TSS, BOD and $NO_3$ as well as performed auto-correlation and cross-correlation analysis (Shim & Tollner, 2014). We used first eight years of data (from January 2001 to December 2008) to build the model and last year of data (from January 2009 to December 2009) to evaluate the model. Within this paper, we showed that the detailed sensitivity analysis of recurrent neural network can allow a better understanding of water quality dynamics of collected runoff and assist in identifying strategies for better management of windrow composting systems.


## 3.1 Introduction

Recurrent neural network (RNN) is a type of artificial neural networks that captures non-linear dynamics of time series. An architectural approach of RNN with embedded memory, Nonlinear Autoregressive model with exogenous inputs (NARX) demonstrated promising qualities for modeling of nonlinear dynamic systems (Xie *et al.*, 2009). Pisoni et al. used NARX neural networks for forecasting peak of ozone concentration in their air pollution study (Pisoni *et al.*, 2009). Jeong et al. applied RNN model for predicting phytoplankton dynamics in a river (Jeong *et al.*, 2001, Jeong *et al.*, 2006). Moreover, Liu et al. adopted RNN for forecasting suspended sediment concentration in a river system (Liu *et al.*, 2013). However, to our knowledge, very few studies applied NARX network for predicting water quality parameters of collected runoff from open windrow composting systems.

Windrow composting systems are alternatives for degrading organic pollutants while preserving opportunities for beneficial reuse. The main advantage of composting pad application is the avoidance of environmental impacts associated with uncontrolled organic pollutants, such as leachate and methane production (Dorahy *et al.*, 2009). Having a composting option is important for certain industrial sectors of interest to state of Georgia. Outreach personnel in Georgia estimate that there are forty to fifty windrow composting facilities in the state. However, open windrow composting systems are exposed to direct rainfall (Dorahy *et al.*, 2009). Typically, windrow composting systems are located in remote areas and runoff cannot be directly treated in a wastewater treatment plant. The main challenge in managing the windrow composting systems is that water quality of pond effluent often exceeds regulatory limits. Pad runoff is highly regulated by U.S. Environmental Protection Agency (EPA) and Georgia Department of Natural Resources (GA DNR) due to pollution potential (Kalaba & Wilson, 2005, Wilson *et al.*, 2004). TSS and BOD regulatory limits are 90 mg/L and 50 mg/L, respectively. Therefore, better management of windrow composting pad is needed in order to meet the regulations for pond effluent.

In our previous study, we showed that a multiple-output neural network demonstrated a better performance than a single-output neural network (Shim & Tollner, 2014). We applied RNN for forecasting runoff water quality constituents such as Total Suspended Solids (TSS), Biological Oxygen Demand (BOD) and $NO_3$(nitrate) levels from windrow composting pad (Shim & Tollner, 2014). We also performed auto-correlation and cross-correlation analysis in order to identify the time series trends and correlations between the variables. One of the challenging research problems in neural network modeling is the black-box nature of this method (Maier & Dandy, 2000). Jeong et al. conducted sensitivity analysis in order to explain phytoplankton dynamics using meteorological, hydrological, physical and chemical parameters (Jeong *et al.*, 2006).

In this paper, we used our previously developed model (Shim & Tollner, 2014) and performed sensitivity analysis using Garson's Algorithm (Garson, 1991), Olden's Method (Olden *et al.*, 2004, Olden & Jackson, 2002), Lek's Profile Method (Lek *et al.*, 1996) and $R^2$-based metric (Giam & Olden, 2015). Additionally, we extended Lek's Profile Method by adopting Sensitivity for Simultaneous Movement of Parameters (SSMP) method from (Kim *et al.*, 2007). SSMP method allowed us extend and better interpret the neural network model as well as describe the relationship between multiple inputs and output variables. As a result, the detailed sensitivity analysis allowed us to gain better insights into the system dynamics behavior and provide recommendations for effective windrow composting pad management.The main objectives of this study were:

(a) extend the application of multiple-input-multiple-output NARX RNN;

(b) to perform detailed sensitivity analysis of RNN in order to identify parameters that are important for predicting water quality constituents;

(c) to better understand the dynamics of the system by ranking the input parameters and quantifying their influence on water quality prediction;

(d) to suggest strategies for effective management of windrow com-posting systems.

## 3.2 Methods

### Windrow composting site characteristics

We used the windrow composting pad located at UGA Bio-conversion center, Athens, GA, USA as our study site (Figure 3.1). The compost material typically shows properties of being hydrophilic as approximately 15 – 20% of rainwater reaching a compost pile is released as surface runoff. The pad with a particle size concentration of 33% sand and 67% gravel sits on top of a tightly compacted Cecil sandy clay loam. The catchment pond has a maximum infiltration rate of less than 3.175 mm/day due to an installed clayed lining (Tollner & Das, 2004). Land application system (LAS) is a method for discharging effluent of a collection pond in order to pre-vent the discharge of organic pollutants. More detailed information of composting pad characteristics can be found in previous studies (Duncan *et al.*, 2013a, Duncan *et al.*, 2013b).

## Observation data and descriptive statistics

We used time series data from nine years (2001 – 2009) of precipitation, air temperature, pond volume, material volume on the pad,total suspended solids (TSS), biological oxygen demand (BOD) and nitrate ($NO_3$) concentration levels of stored runoff in the collection pond. The Georgia Environmental Protection Division mandates collection, sample handling and analysis. Daily data of precipitation and temperature represent observations from Georgia Automated Environmental Monitoring Network (GA AEMN) site weather station. Weather data follows Quality Control of National Oceanic and Atmospheric administration (NOAA). Daily data of pond volume were obtained from gauge readings of pond stage. A surveyor employed by the Grounds Department at UGA collected monthly data of material volume on the pad using photogrammetric techniques. For water quality parameters data, monthly water samples of 1 L were collected in Nalgene containers at the site and sent to the UGA Feed and Environmental Water laboratory for analysis. The samples were analyzed according to Standard Methods for the Examination of Water and Wastewater as well as Methods for Chemical Analysis of Water and Wastes. Water samples were analyzed for TSS and BOD every month and for $NO_3$ each quarter. Detailed description of data collection and water quality analysis can be found in the previous study of (Shim & Tollner, 2014).

Supplementary data in Appendix A (Figures A1-A7) represent time series data of precipitation, air temperature, pond volume, material volume on the compost pad, TSS, BOD and $NO_3$ at monthly resolution. Additionally, each figure contains yearly and monthly box-plots along with a histogram. For Figures A1-A7, daily precipitation was monthly cumulated, daily air temperature and daily pond volume were monthly averaged. Additionally, quarterly nitrate data were interpolated to obtain uniform monthly data set.

Figure A1 demonstrates that high precipitation levels usually occur during winter months whereas extreme storm events tend to appear during summer. As it is expected, air temperature has a seasonal pattern with the highest temperature usually observed in August (Figure A2). From Figure A3 we observe that low pond volume level occurs during April-June period (when the precipitation level is low as well). We notice an increasing trend in waste material volume on the pad from 2001 to 2007 (Figure A4). From the monthly box-plot in Figure A4, we see that waste volume reaches its low level during March-June period.

Figures A5 and A6 show regulatory limits for TSS and BOD concentration. We observe that TSS concentration often exceeds the regulatory threshold (90 mg/L) especially during period from May to September whereas high BOD concentration occurs during April-September period. On the other hand, $NO_3$ reaches the lowest concentration levels during a period from June until September (Figure A7).

Table 3.1 shows descriptive statistics of both input and output variables. Because we need to normalize our variables between 0 and 1, this information is helpful in interpreting results of neural network modeling and sensitivity analysis.

## Artificial neural network modeling

Artificial Neural Networks (ANNs) are computational models that represent collection of neurons based on the structure and operation of the biological nervous system. Figure 3.2 shows an ANN with three layers. The first input layer has a set of neurons that receive a signal input ($x$); the second hidden layer contains hidden neurons ($u$) that receive information from neurons of input layer; the third output layer consists of output neurons ($v$) that represent the response of the network ($y_p$).

The layers are connected by the connection weights ($a$ and $b$). For processing information, one or several hidden layers can be used. The number of neurons in the input ($I$) and output ($K$) layers corresponds to the number of input and output variables, respectively. Through experimentation, one may estimate the optimal number of neurons in the hidden layer ($J$).

**Recurrent neural network (RNN) models structure overview**

ANNs can have different architectures that lead to various types of neural networks. A multilayer perceptron (MLP) or feedforward network is a neural network that includes layers of parallel perceptrons. In MLP neural network, only forward connection from one neuron in a layer to neurons of the next layer is allowed (Eberhart & Shi, 2007). On the other hand, recurrent neural network (RNN) can have feedback connection from either the output layer or hidden layer to the input layer. This allows RNN to have memory and to be trained to learn time-varying patterns (Araghinejad, 2014). Figure B1 in Appendix B represents RNN that contains several time-delay operators that represent dimension of time.

**NN training procedure and experimental setup**

In this section, we describe the water quality recurrent neural network model developed in our previous study (Shim & Tollner, 2014). Because for training purposes of ANN, large dimension of data is required (Ssegane *et al.*, 2009); we used daily time series data for precipitation, air temperature and pond volume variables. We used a step function to transform monthly data for material volume, TSS and BOD as well as quarterly data for $NO_3$ to daily resolution time series by making an assumption that these values do not dramatically change over a month. The input variables were precipitation, air temperature, pond volume and waste material volume; the output variables were TSS, BOD and $NO_3$ (Table 3.1).

For building the neural network model, we used data from 2001 to 2008 where each variable's sample size was 2922. We divided the data set of each variable into three subsets: training, validation and testing data (Table 3.2). Additionally, we studied effect of number of hidden nodes on neural network performance by changing number of hidden nodes from 5 to 20.

During training or learning process, initially random generated weights were corrected by comparing the ANN predicted value to target value using training iteration of 1000 according to the following steps:

1. reading values of neurons in input layer;

2. calculating input (or activation) function and transfer function for neurons in hidden layer using Equations (3.1) and (3.2), respectively;

$$u_j = a_{0j} + \sum_{i=1}^{I} x_i a_{ij} \tag{3.1}$$

$$z_j = g(u_j) = \frac{1}{1 + e^{-u_j}} \tag{3.2}$$

   where $u_j$ is the $j^{th}$ neuron in the hidden layer, $a_{0j}$ is the bias weight for the hidden layer, $x_i$ is $i^{th}$ input variable, $a_{ij}$ is the connection weight from the $i^{th}$ input neuron to $j^{th}$ hidden neuron, $g(u_j)$ is the transfer function yielding the output $z_j$.

3. calculating input (or activation) function and transfer function for neurons in output layer using Equations (3.3) and (3.4), respectively;

$$v_k = b_{0j} + \sum_{i=1}^{I} z_i b_{ij} \tag{3.3}$$

$$y_{p_k} = g(v_k) = \frac{1}{1 + e^{-v_k}} \tag{3.4}$$

   where $v_k$ is the $k^{th}$ neuron in the output layer, $b_{0_j}$ is the bias weight for the output layer, $b^{ij}$ is the connection weight from the $i^{th}$ hidden neuron to $j^{th}$ output neuron, $g(v_k)$ is the transfer function yielding the $k^{th}$ predicted output $y_p$.

4. calculating error between ANN predicted value $(y_p)$ and target value $(y_t)$ using Equation (3.5);

$$E = \frac{1}{2}(y_p - y_t)^2 \tag{3.5}$$

5. calculating weight change for optimal weight adjustment in order to minimize the above error.

Steps 1 – 4 describe feedforward method wherein step 5 represents back propagation method (using the Levenbergh-Marquardt algorithm). The above training procedure was repeated until no further improvement of ANN could be achieved using the validation set. We used the error calculated for validation set using Equation (3.5) as a stopping criteria. We stopped the training procedure when for hundred iterations the validation error increased (Beale *et al.*, 2012) to avoid overfitting. The testing set provided an independent measure of network performance during and after training (Ssegane *et al.*, 2009).

For evaluating the neural network model, we used data from 2009 (sample size of each variable was 365). In order to find the best model performance, we compared $R^2$ and RMSE values calculated using Equations (3.6) and (3.7), respectively.

We used the MatLab Neural Network Time Series Toolbox (Beale *et al.*, 2012) for building recurrent neural network and evaluating its performance.

$$R^2 = \frac{\sum_{k=1}^{K}(y_{t_k} - \bar{y_t})^2}{\sum_{k=1}^{K}(y_{p_k} - \bar{y_p})^2} \tag{3.6}$$

$$RMSE = \sqrt{\frac{1}{k-1}\sum_{k=1}^{K}(y_{t_k} - \bar{y_{p_k}})^2} \tag{3.7}$$

## Sensitivity analysis of neural network

The main goal of sensitivity analysis is to obtain knowledge of the relationship between independent and dependent variables in a model. We achieved it by ranking the input parameters and quantifying their influence on output parameters prediction. For sensitivity analysis, we used "NeuralNetTools" package in R (Beck, 2015). In order to interpret the neural network model and describe the connection between the variables, we focused on the following methods: Garson's Algorithm (Garson, 1991), Olden's Method (Olden & Jackson, 2002), Lek's Profile Method (Lek *et al.*, 1996), $R^2$-based variable importance metric (Giam & Olden, 2015) and Sensitivity for Simultaneous Movement of Parameters (SSMP) method (Kim *et al.*, 2007). An example of detailed calculation of weight importance performed by Garson's and Olden's algorithms can be found in Figure B2 in Appendix B (Olden *et al.*, 2004). Here, we provide a brief description of the five methods.

### Garson's algorithm

The method proposed by (Garson, 1991) is based on partitioning neural network connection weights. Similar to regression coefficients, connection weights are used in order to define relationship between variables. In order to quantify influence of inputs on the output, the algorithm reports relative importance of each input variable expressed as a percentage. (Goh, 1995) described a detailed procedure of calculating relative importance of the inputs. The calculation includes hidden-input layer and hidden-output layer connection weights associated with output neuron and each input neuron. Higher value of connection weight indicates stronger importance of corresponding input variable. The algorithm uses absolute values of connection weights and it can be applied only to neural networks with one output variable.

### Olden's algorithm

(Olden *et al.*, 2004) introduced another method for obtaining relative importance of the input variables. As the Garson's algorithm, Olden's algorithm (also called as "Connection Weights" Approach by (Olden & Jackson, 2002)) involves connection weights between layers of neural network but it uses raw values of connection weights instead of absolute values of connection weights. Therefore, it accounts for positive or negative signs of the weights by calculating the sum of the product of the raw input-hidden layer and hidden-output layer connection weights between output and each input neuron. Negative sign of connection weight corresponds to decrease in output value wherein positive sign of connection weight suggests increase in output value (Olden & Jackson, 2002). The Olden's algorithm can be applied to neural networks with multiple outputs.

### Lek's profile method

Lek's profile method performs sensitivity analysis by looking at response of neural network to input variables (Lek *et al.*, 1996, Gevrey *et al.*, 2003). This method produces profile (or contribution) plots of each output variable with respect to a range of one input variable wherein the rest of input variables are held constant at their $0^{th}$, $25^{th}$, $50^{th}$, $75^{th}$ and $100^{th}$ percentile. Afterwards, the method repeats this process for each input variable. As a result, it constructs response curves according to the change of the input variables (Gevrey *et al.*, 2003). Table 3.1 provides information regarding range of input and output variables as well as five different percentile values.

### $R^2$-based metric

$R^2$-based metric is a new method of evaluating importance of input variables using $R^2$ of ANN and permutation of each input in turn (Giam & Olden, 2015). This method computes the permutational relative variable importance ($pRVI$) using Equation 3.8:

$$pRVI = R^2_{obs} - \bar{R}^2_{perm} \tag{3.8}$$

where $R^2_{obs}$ is the $R^2$ of the ANN model and $\bar{R}^2_{perm}$ is the mean of $R^2$ of the ANNs fitted to modified datasets where each of the input variables is permuted in turn. The higher $pRVI$ value associated with an input variable, the more important the variable. This method is an alternative tool to Garson's and Olden's Algorithm and it demonstrated more accurate performance in measuring variable performance (Giam & Olden, 2015).

**Sensitivity for simultaneous movement of parameters (SSMP)**

Sensitivity for simultaneous movement of parameters (SSMP) method examines response of output variables to the changes in multiple inputs at a time (Kim *et al.*, 2007). For this study, we selected the two most important input variables identified based on methods from sensitivity analysis (described in Section 3.2). Using this method, we generated $3D$ plots to visualize how each output variable behaves in response to simultaneous changes of the two inputs.

## 3.3 Results

Based on our previous study (Shim & Tollner, 2014), the recurrent neural network demonstrated the best performance with training/validation/testing ratio = 70/15/15, number of hidden nodes = 10. The RNN achieved $R^2$ of 0.99, 0.98, 0.99 and RMSE of ± 3.49, ± 3.79, ± 0.07 for TSS, BOD and Nitrate predictions, respectively. Therefore, we used this architecture of the neural network with four inputs and three outputs in order to perform sensitivity analysis according to Garson's algorithm, Olden's algorithm, Lek's profile method and $R^2$-based metric.

## Sensitivity analysis for total suspended solids

### Garson's and Olden's Algorithms

According to Garson's algorithm, for Total Suspended Solids (TSS) prediction, Temperature and Waste Volume are the most important input variables followed by Pond Volume and Precipitation (Figure 3.3A). However, based on Olden's algorithm, Pond Volume is the most important input variable for predicting TSS concentration (Figure 3.3B) and its importance value has a negative sign which suggests that pond volume reduces TSS concentration. On the other hand, Waste Volume, Temperature and Precipitation have positive signs which indicate that they contribute to TSS concentration increase.

### Lek's profile method

Lek's profile method demonstrates that Pond Volume contributes the most to TSS concentration when the other three input variables are held at their minimum (Figure 3.3C) suggesting that it is the most important factor during dry winter months when waste volume is minimum. Additionally, TSS has an increasing trend as Precipitation increases when Pond Volume, Temperature and Waste Volume are held at their minimum. Therefore, rain events contribute to TSS levels in winter when Pond Volume and Waste Volume are minimal. On the other hand, during summer months when Pond Volume and Waste Volume are maximum, Precipitation contributes the most at low values (dry weather) and high values (extreme storm events). Moreover, TSS constantly increases across the range of Temperature when the other three inputs are maximum which means that high temperatures lead to TSS increase when Pond Volume, Waste Volume are the highest during storm events. When Precipitation, Pond Volume and Waste Volume are at their $0^{th}$, $25^{th}$, $50^{th}$ and $75^{th}$ percentiles Temperature affects TSS concentration at $25^{th}$ percentile. Furthermore, TSS concentration linearly increases over the range of Waste Volume and changes its magnitude when we hold Pond Volume, Precipitation and Temperature at five different percentile values.

### $R^2$-based metric

$R^2$-based metric identified the two most important input variables as Pond Volume and Waste Volume for TSS. Variable importance ranking identified by values of permutational relative variable importance ($pRVI$) calculated using Equation 3.8. Variable importance was defined as $pRVI_{PondVolume} > pRVI_{WasteVolume} > pRVI_{AirTemperature} > pRVI_{Precipitation}$.

## Sensitivity analysis for biological oxygen demand

### Garson's and Olden's Algorithms

Using results from Garson's algorithm for Biological Oxygen Demand (BOD), we see that Temperature and Waste Volume are the most important input variables followed by Pond Volume and Precipitation (Figure 3.4A). Olden's algorithm results also suggest that the most important input variables for predicting BOD concentration are Temperature and Waste Volume (Figure 3.4B). It also demonstrates that Temperature and Pond Volume increase BOD concentration whereas Waste Volume and Precipitation decrease BOD concentration in the pond.

### Lek's profile method

According to Lek's profile method, Pond Volume contributes minimally across its range to BOD (Figure 3.4C) which confirms the results from Olden's algorithm (Figure 3.4B). BOD concentration increases across Precipitation range when the other three input variables are at $0^{th}$, $75^{th}$ and $100^{th}$ percentiles (Figure 3.4C). This suggests that high BOD concentration occurs when it rains a lot during summer months and when both Pond Volume and Waste Volume are above average levels. An average temperature contributes to BOD during dry weather when Pond Volume and Waste Volume are minimum. When the three input variables are at their maximum, BOD constantly increases across the range of Temperature (Figure 3.4C) suggesting highest BOD levels in summer months when storm events as well as maximum levels of pond and waste volume occur. BOD and

TSS have similar trend over the range of Waste Volume as Waste Volume contributes to linearly increase of BOD and TSS concentration levels (Figure 3.3C and 3.4C).

### $R^2$-based metric

Temperature and Waste Volume are the two most important inputs for BOD as indicated by $R^2$-based metric. The method suggested that importance of the input variables follows in order: $pRVI_{AirTemperature} > pRVI_{WasteVolume} > pRVI_{Precipitation} > pRVI_{PondVolume}$.

## Sensitivity analysis for nitrate ($NO_3$)

### Garson's and Olden's Algorithms

Based on Garson's algorithm results for Nitrate ($NO_3$) response, we see that the most important variables are Waste Volume and Pond Volume followed by Temperature and Precipitation (Figure 3.5A). Olden's algorithm results also show that the most important input variable for Nitrate is Waste Volume as it contributes to $NO_3$ the most whereas Pond volume decreases $NO_3$ concentration (Figure 3.5B).

### Lek's profile method

By analyzing graphs from Lek's profile method, $NO_3$ response curves have an increasing trend across the range of Pond Volume when we hold Precipitation, Temperature and Waste Volume variables constant at their $0^{th}$ and $25^{th}$ percentile values (Figure 3.5C). If we look at contribution plots of Nitrate with respect to the range of Precipitation, we notice that Precipitation contributes decreasingly at increasing values. The highest Nitrate concentration corresponds to dry summer months when both Pond Volume and Waste Volume levels are high. During storm events when precipitation falls on a full pond and waste volume on the compost pad is high, $NO_3$ concentration follows a U-shape non-linear pattern. Nitrate concentration is high during both low and high temperatures and it is minimum with intermediate temperatures (Figure 3.5C). $NO_3$ concentration increases till Waste Volume reaches its $50^{th}$ percentile and then it starts to decrease when the other three variables are at $25^{th}$, $50^{th}$ and $75^{th}$ percentile.

#### $R^2$-based metric

According to $R^2$-based metric method, $pRVI_{WasteVolume} > pRVI_{PondVolume} > pRVI_{Precipitation} > pRVI_{AirTemperature}$. Therefore, waste volume and pond volume are the most important input parameters for $NO_3$.

## 3.4 Discussion

### Garson's Algorithm, Olden's Algorithm and $R^2$-based metric

Sensitivity testing with neural networks is a continuing work. Rather than claiming that one technique is better than another, we chose to base our operational recommendations on common trends among the sensitivity methods. Table 3.3 provides comparison of the results from the three methods that identified the most important and the least important input parameters coded as *P* for Precipitation, *T* for Temperature, *PV* for Pond Volume and *WV* for Waste Volume. Our results suggest that Olden's method gives different results when compared to Garson's method. This is because Olden's method accounts for variable's relative importance based on magnitude and sign of connection weights wherein Garson's method considers only absolute values of connection weights. It is also consistent with the results reported by other researchers (Gevrey *et al.*, 2003, Nourani & Fard, 2012, Olden *et al.*, 2004). $R^2$-based metric results were similar to the results obtained from Olden's method which is also consistent with (Giam & Olden, 2015).

## Lek's profile and SSMP methods

Additionally, Lek's profile method not only classifies the input variables by relative importance but also describes how these inputs contribute to the output (visualized in contribution plots). The trends and results from Lek's profile method reflect that the method also includes interaction between the variables. As an extension of the Lek's profile method, Sensitivity for Simultaneous Movement of Parameters (SSMP) method visualizes changes in output variables with respect to the two most important variables (identified using $R^2$-based metric) in $3D$ plots (Figure 3.6). These $3D$ plots were helpful in gaining more insights into dynamic behavior of runoff water quality parameters of the windrow composting system.

## Recommendations for windrow composting system management

This study makes a case for modifying management guidelines for regulators to correct pond storage and spraying system management based on our results. Keeping waste volume at low levels and pond volume at medium levels is preferred if the operators of windrow composting system to maintain TSS and $NO_3$ levels (Figures 3.6A and 3.6C). Additionally, operators can prevent high BOD levels by keeping waste volume at low levels during winter and summer months (Figure 3.6B).

As our results suggest, air temperature and pond volume (followed by waste volume) are the most significant input variables contributing to the outputs. Due to high precipitation during summer months, windrow composting pad operators usually try to maintain the pond volume at lower levels during spring and summer in order to avoid overtopping the pond. The resulting shallow pond experiences elevated temperatures, which leads to excessive TSS concentration and BOD levels. Additionally, acceleration of nitrification processes occurs at temperature above $17°C$ as composting pad experience high temperatures. This results in increase of BOD and nitrate concentration levels.

The need to maintain pond capacity for extreme rainfall events necessitates the shallow pond.

One practical response is to spray the water from the pond on the composting pad instead of discharging the effluent during summer and early fall months. Relying solely on spraying effluent back on the compost pad may have other unintended consequences due to possible buildups of nutrients and small soluble solids on the pad. Having a pond with sufficient capacity to maintain a four to six foot depth of operational storage while having additional storage for an extreme event storm (a 25 or 50 year storm) would also be an acceptable remedy in that deeper ponds tend to remain cooler, which suppresses biochemical activity in the pond.

## 3.5 Conclusion

In this study, we extended application of recurrent neural network (RNN), performed sensitivity analysis of RNN and achieved abetter understanding of system dynamics. We also demonstrated the significance of the detailed sensitivity analysis in order to overcome the black-box nature of the artificial neural network and to advance our knowledge about windrow composting system dynamics in the absence of detailed relationships among the factors. The sensitivity analysis brought into clearer focus the effect of heating in shallow ponds and the resulting increase in TSS and BOD, leading to both increased pond capacity needs and some alternate disposal approaches in summer and early fall months. The modeling approach of recurrent neural network can be further applied in other studies of dynamic ecological systems modeling because it generalizes in a straightforward manner to nearly any scenario.

## Acknowledgements

We would like to thank Dr. Nathan D. Melear and research staff at bioconversion laboratory of College of Engineering, University of Georgia (UGA) for their assistance in data collection and management. We thank UGA Feed and Environmental Water Lab for performing water quality analysis. We also thank the two anonymous referees and the associate editor for their valuable comments.

Table 3.1: Descriptive Statistics of input and output variables.

| Variable | | Mean | SD | Percentile | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | $0^{th}$ | $25^{th}$ | $50^{th}$ | $75^{th}$ | $100^{th}$ |
| Input | Precipitation (mm) | 2.973 | 9.95 | 0 | 0 | 0 | 0 | 142.24 |
| | Pond Volume ($m^3$) | 737,663 | 475,409 | 87,882 | 328,131 | 719,436 | 971,734 | 2,799,697 |
| | Waste Volume ($m^3$) | 4507 | 1262.60 | 1775 | 3677 | 4358 | 5367 | 7532 |
| | Temperature (°C) | 16.75 | 8.01 | -10 | 10 | 18 | 24 | 32 |
| Output | TSS (mg/L) | 50.41 | 56.26 | 1 | 10.17 | 50.41 | 72 | 290 |
| | BOD (mg/L) | 29.67 | 25.34 | 2 | 13 | 21.09 | 37.47 | 124 |
| | Nitrate (mg/L) | 0.84 | 0.82 | 0.02 | 0.18 | 0.5 | 1.3 | 3.5 |

Table 3.2: Training, Validation and Testing Data used for building RNN.

| | Data Percentage% (n = Sample Size) | | |
|---|---|---|---|
| | Training | Validation | Testing |
| 1 | 50% (n = 1462) | 25% (n = 730) | 25% (n = 730) |
| 2 | 60% (n = 1754) | 20% (n = 584) | 20% (n = 584) |
| 3 | 70% (n = 2046) | 15% (n = 438) | 15% (n = 438) |
| 4 | 80% (n = 2338) | 10% (n = 292) | 10% (n = 292) |

Table 3.3: Variable Importance based on the Garson's, Olden's and $R^2$-based metric methods ($P$ is Precipitation, $T$ is Temperature, $PV$ is Pond Volume and $WV$ is Waste Volume).

| Method | Variable Importance for TSS/BOD/NO3 | |
|---|---|---|
| | Most Important Variable(s) | Least Important Variable(s) |
| Garson's Algorithm | $T$ and $WV$ / $WV$ / $PV$ and $WV$ | $P$ |
| Olden's Algorithm | $PV$ / $T$ / $WV$ | $P$ |
| $R^2$-based metric (pRVI) | $PV$ / $T$ / $WV$ | $P$ |

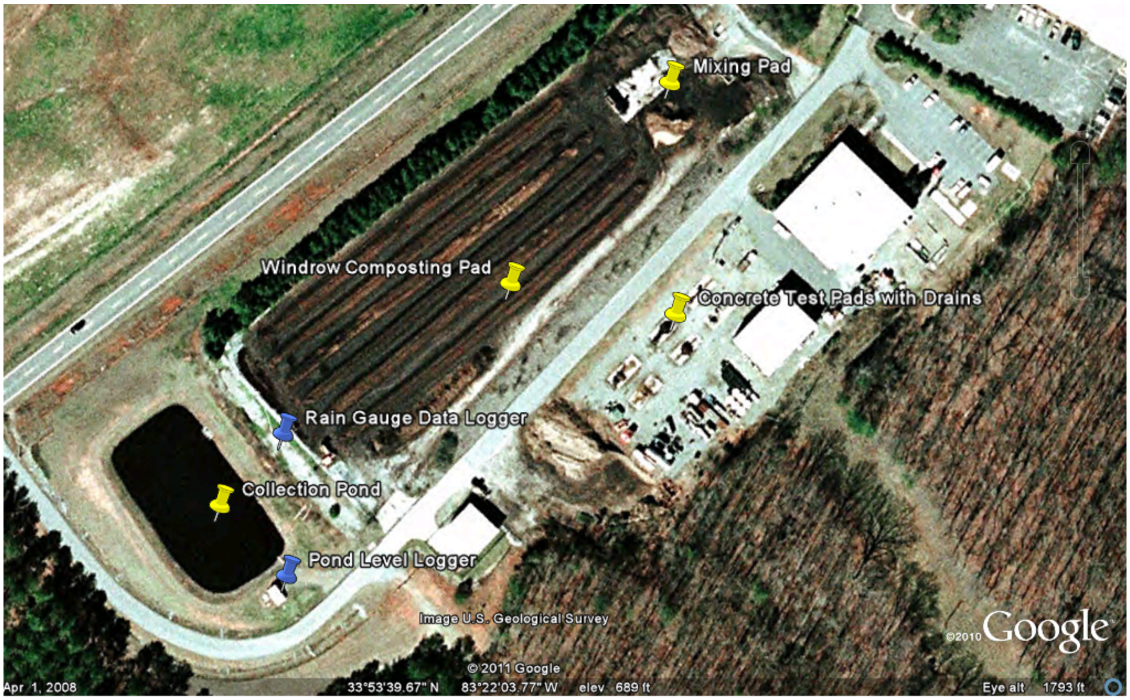



Figure 3.1: Google Earth (© 2010) overhead view of the windrow composting pad at the UGA Bioconversion Center (Duncan *et al.*, 2013b).

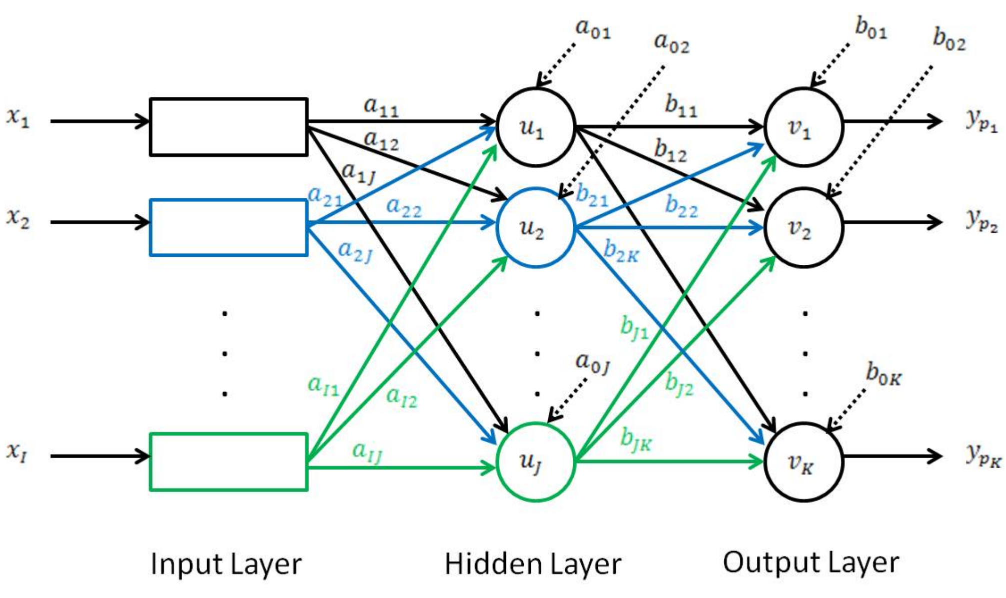



Figure 3.2: Schematic representation of a three layer neural network.

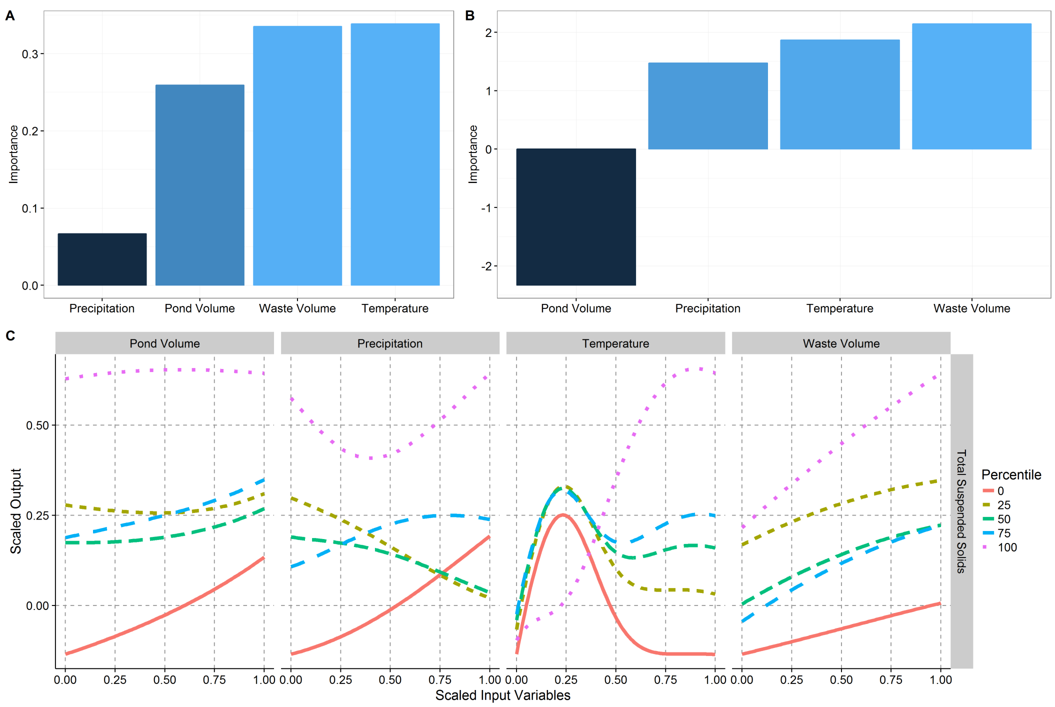



Figure 3.3: Sensitivity Analysis for Output variable: Total Suspended Solids (A) Garson's algorithm (B) Olden's algorithm (C) Lek's profile method.

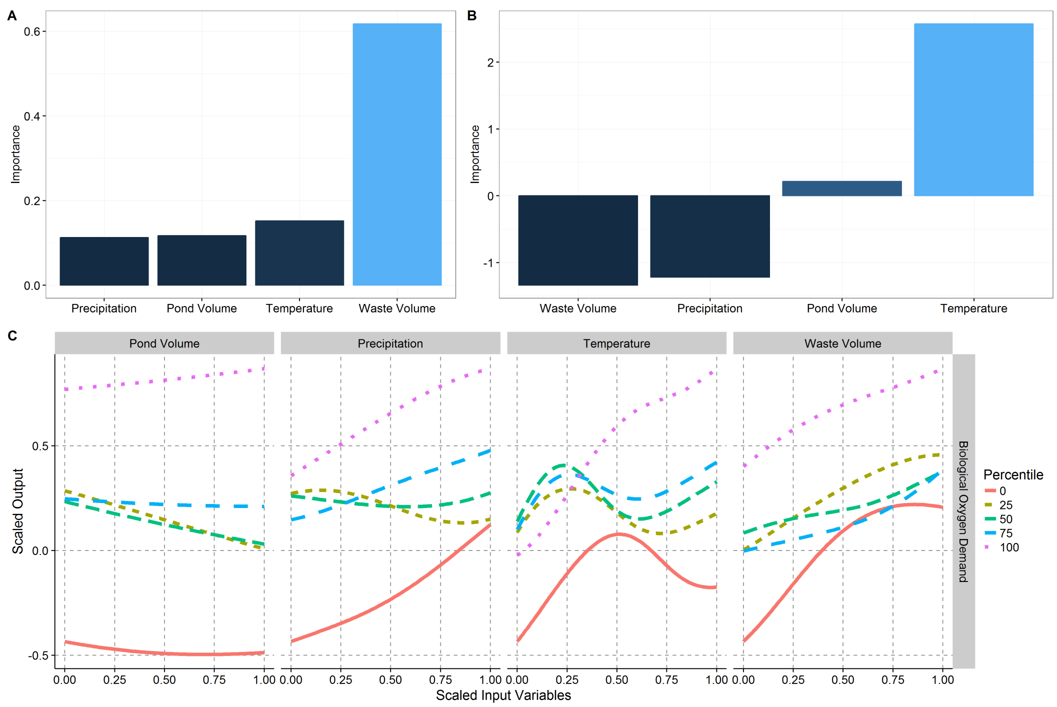



Figure 3.4: Sensitivity Analysis for Output variable: Biological Oxygen Demand (A) Garson's algorithm (B) Olden's algorithm (C) Lek's profile method.

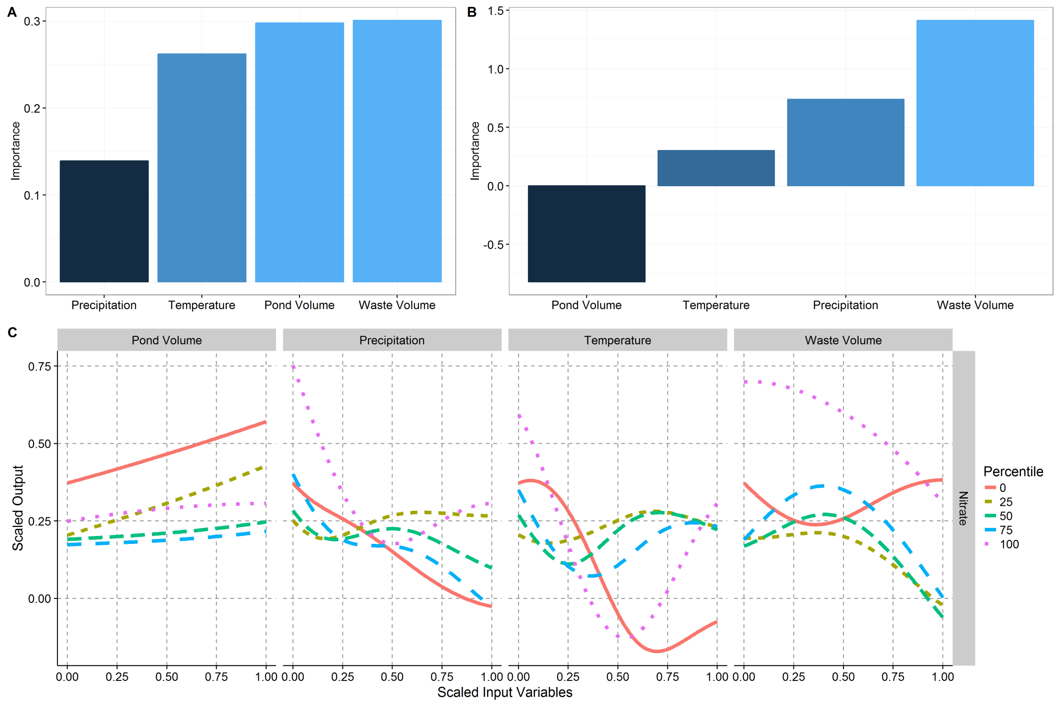



Figure 3.5: Sensitivity Analysis for Output variable: Nitrate (A) Garson's algorithm (B) Olden's algorithm (C) Lek's profile method.

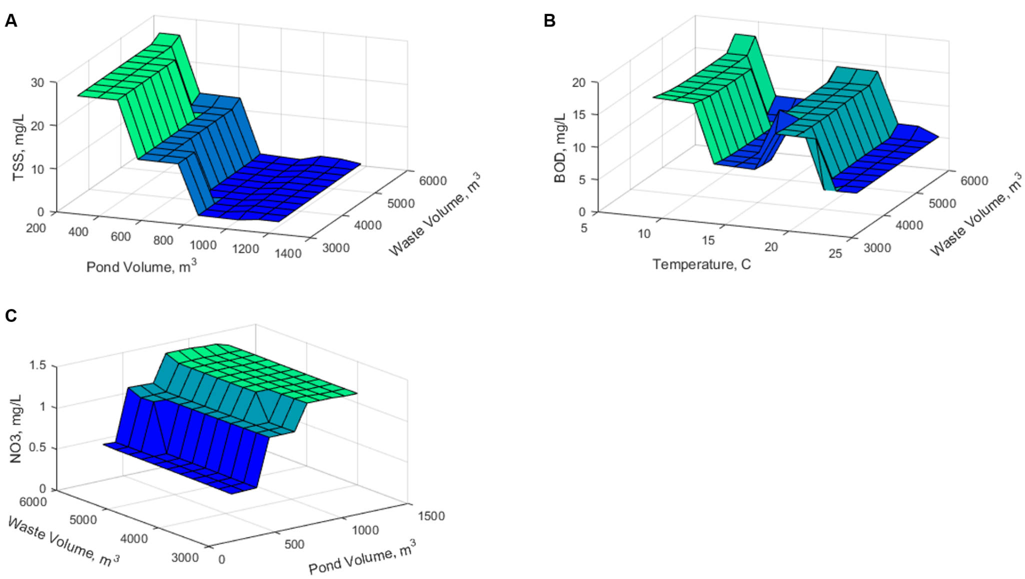



Figure 3.6: Sensitivity Analysis using SSMP method for Output: (A) Total Suspended Solids (B) Biological Oxygen Demand (C) Nitrate.

# Chapter 4

# Habitat provision associated with environmental flows in the Middle Oconee river[1]

[1] N. V. Bhattacharjee, J. R. Willis, S. K. McKay and E. W. Tollner. Submitted to *US Army Engineer Research and Development Center*, 12/17/2016

# Abstract

A comprehensive analysis of environmental flow schemes represents a key step in ensuring adequate water availability to meet increasing human needs while minimizing adverse impacts on aquatic ecosystems. This study investigates how several environmental flow methods affect habitat provision in the Middle Oconee River near Athens, Georgia. Historical discharge data are coupled with water withdrawal simulations for each of the environmental flow types to examine trade-offs between ecological and social outcomes (i.e., habitat provision and water withdrawal, respectively). Hydraulic models are applied to translate hydrologic simulations into habitat suitability for three generic habitat types: shallow-fast, deep-fast, and shallow-slow. The availability and distribution of habitats are then analyzed with respect to increasing water withdrawal rates. Finally, we compare the utility of deterministic modeling approaches based on long-term average conditions relative to stochastic modeling approaches using frequency-weighted outcomes. The analytical methodology and approach set forth in this work may be easily adapted to inform environmental flow analyses at other study sites.

## 4.1 Introduction

Human society, culture, and economy are vitally dependent upon freshwater ecosystems; however, they are becoming increasingly compromised (Strayer & Dudgeon, 2010). For instance, human uses capture more than 50% of available freshwater runoff, and upwards of $1,000,000$ dams fragment river systems globally (Jackson *et al.*, 2001). These factors, in addition to many others, have led to rivers becoming the earth's most damaged ecosystems, losing species at greater rates than terrestrial and marine ecosystems (Dudgeon *et al.*, 2006). Accordingly, sustainable strategies for river management are needed which balance aquatic ecosystem integrity and human livelihoods that rely on them, and significant research and management has been devoted to the subject. Environmental flows provide at least a partial solution to these freshwater management challenges.

In the 2007 “Brisbane Declaration”, over 750 scientists, engineers, and lawmakers from around the world defined environmental flows as “the quantity, timing, and quality of water flows required to sustain freshwater and estuarine ecosystems and the human livelihoods and well-being that depend on these ecosystems” (Declaration, 2007).

The purpose of this work is two-fold. First, we demonstrate a set of methods for comparing simple environmental flow alternatives and developing flow thresholds, which applies a common approach to habitat analyses. Second, we examine the role of variability in setting flow management thresholds by comparing environmental flow recommendations developed using “typical” river levels (i.e., long-term averages) with those developed using a frequency-weighted approach (i.e., effectiveness analysis). The analysis and findings of this research focus on a case study in the Middle Oconee River near Athens, Georgia, but the techniques applied are transferrable to applications elsewhere.

## 4.2 Methods

Here, we examine the ecological effects of different environmental flow thresholds for the Middle Oconee River. First, we simulate hydrologic alteration associated with three different environmental flow approaches. Second, we construct a hydraulic model to translate changes in discharge into changes in velocity and depth regimes (ecologically relevant habitat variables). Third, we apply generalized habitat criteria to quantify changes in the general habitat composition in the Middle Oconee River as a result of municipal water withdrawal. Figure 4.1 visualizes the framework of this research.

### Middle Oconee River, Georgia

As a result of growing human population and economic development in northeast Georgia, surface water withdrawal is projected to increase (NGRC, 2011, UORWPC, 2011). The Middle Oconee

River is a sixth-order tributary of the Altamaha River located in northeast Georgia. The study area encompasses a small reach near Ben Burton Park in Athens, Georgia, which is approximately 365 meters long and 45 − 60 meters wide. Rapid development in northeast Georgia has resulted in increased municipal water demand, and as a response, Bear Creek Reservoir was constructed in 2002 to meet municipal water needs of four surrounding counties. Bear Creek Reservoir is an off-channel reservoir located on a tributary but filled with water from the Middle Oconee River. Since 1938, the U.S. Geological Survey (USGS) has operated a long-term streamflow monitoring station nearby (Gage number 02217500, Figure 4.2). The reservoir is permitted to withdraw a maximum of 60 million gallons per day (MGD; Georgia EPD Permit Number 078 − 0304 − 05) subject to meeting minimum flow criteria. Currently the reservoir typically withdraws less than 20 MGD (Campana *et al.*, 2012), but the permitted rate represents a substantial portion of river discharge (60 MGD = 2.63 L/s = 92.8 cfs), particularly during the late summer months when flow rates are lowest (September mean = 6.7 L/s or 237 cfs). Thus, this system provides unique long-term flow data set with minimal flow alteration as well as an opportunity to examine the effects of flow management actions for a pump-storage configuration, which are becoming more common in the region. Figure 4.2 also includes annual $7Q10$ and monthly $7Q10$ estimates marked as blue dashed line and solid green line, respectively. (McKay, 2014).

Based on the historical data 1938 − 1997 (before Bear Creek Reservoir construction), daily mean streamflow fell below the annual $7Q10$ of 1 L/s or 37 cfs during July - October period. There were 113 days when the flow was recorded to be less than the annual $7Q10$ . The histogram of low flow distribution during 1938 − 1997 is provided in Figure 4.3.

## Flow Management

Daily discharge records are continuously available from 1938 to present. However, we use a period of analysis from 1938 − 1997, which represents a minimally altered flow regime and is approximately the data available at the time of permit application (Figure 4.2). Four scenarios of municipal water withdrawal and environmental flow requirements were simulated (304 total simulations):

1. Unaltered: A reference condition without withdrawal defined the best attainable ecological condition and served as a point of relative comparison for other scenarios.

2. Annual minimum flow (AMF): This method assigns a single, year-round flow threshold below which water may not be withdrawn. The minimum flow threshold was varied from 0 to $1,000$ cfs by 10 cfs increments to assess the influence of minimum flow magnitude on ecological conditions.

3. Monthly minimum flow (MMF): This method assigns a monthly-varied flow threshold below which water may not be withdrawn, which incorporates elements of water availability not captured in annual minimum flows. Flow thresholds were varied in 101 intervals from the minimum observed monthly-averaged flow to the maximum observed monthly-averaged flow for the 60-year record for each of the 12 months.

4. Percent of flow (POF): This method withdraws a specified percentage of the unaltered discharge, which was varied from 0 to 50% by 0.5%.

   For each simulation, the unaltered hydrograph was modified for the entire 60-year observational period (i.e., 1938 – 1997). Water was abstracted at a maximum rate of 60 MGD in accordance with existing pump capacity. Although previously acknowledged as operational constraints (Vogel *et al.*, 2007), neither reservoir volume limitations nor increased water treatment costs due to turbidity of high flows were included in this analysis. Each simulation produced a 60-year record of daily river discharge and daily water withdrawal, which are subsequently used in habitat trade-off analyses. Numerical models were developed in the R statistical software (version 3.2.5), and code is available in Appendix C.

## Hydraulic Modeling

While hydrologic alteration is a common surrogate for ecological integrity (McKay, 2015), habitat analyses require that hydrologic change be converted into hydraulic variables (e.g., velocity and depth), which are often more ecologically relevant. Here, the USACE Hydrologic Engineering Center's River Analysis System (HEC-RAS version 4.1.0) is applied to assess channel hydraulics along with the accompanying HEC-GeoRAS (version 10.1), which facilitates geospatially explicit analyses in ArcGIS® (version 10.2.2) (Brunner, 2001).

Topographic and bathymetric data provide a crucial basis for modeling river hydraulics. We characterized the terrain at the study site by merging three topographic data sets. First, high-resolution Light Detection and Ranging (LiDAR) data was obtained from local government (Personal Communication, Mary R. Martin, GIS Administrator, Athens-Clarke County Planning Department), which provided a $1m \times 1m$ gridded digital elevation model for the bank and floodplain zones. Second, eleven cross-sectional surveys were collected at the site in June - July 2013 to better characterize underwater bathymetry, a common gap in LiDAR measurements (Personal communication, Dr. Bruce Pruitt, USACE Environmental Laboratory). Third, ten additional cross-sections were collected in November − December 2015 to fill key gaps in surveyed cross-sections (Figure 4.4). Four of the cross-sections collected in 2013 and 2015 overlapped and were used to check for consistency between the different survey teams. Both surveys included points in the bank and floodplain zone to assist in merging with the LiDAR data. The inverse distance weighting (IDW) tool in ArcGIS® (version 10.2.2) was used to interpolate between elevation data gaps to form a gridded elevation layer for the LiDAR and two surveys. We assumed the bathymetry of the stream to be unchanging.

All three data sets were stitched together with LiDAR data used for floodplain zones (i.e., above the top of bank elevation of 570 ft) and survey data for channel and bank zones (i.e., below the top of bank elevation of 570 ft). Ultimately, this process generated an integrated digital elevation model of the topography and bathymetry of the study site (Figure 4.5). The combined raster data were converted to a triangular irregular network (TIN) as the primary input to HEC-RAS.

In addition to terrain, HEC-RAS requires user-inputs related to flow paths, channel roughness, and channel slope. Flow paths were demarcated using HEC-GeoRAS. Following standard convention, floodplain flow paths were estimated as the center of mass between the top of the bank and the extent of the floodplain (roughly $\frac{1}{3}$ of the distance from the banks and $\frac{2}{3}$ from the floodplain extent). Channel roughness (i.e., Manning's $n$) was estimated through an iterative, pseudo-calibration process. Manning's $n$ was predicted from standard tabulated values for the channel and floodplain environments (Chow, 1959).

Roughness values were then iteratively adjusted based on two sets of observed water surfaces:

1. water surface elevations during survey periods;
2. fresh alluvial sand deposits accumulated following a high flow in December 2015.

This process resulted in four distinct values of Manning's $n$: 0.065 for the open, moderately vegetated left floodplain, 0.070 for the more densely vegetated right floodplain, 0.025 for sandy portions of the channel, and 0.040 for rocky or "shoaly" portions of the channel. To estimate channel slope, thalweg measurements in each cross-section (i.e., the deepest point) were coupled with longitudinal distance downstream. The slope of the best fit, linear regression of these data were used as the slope value in simulations, 2.45 ‰. These inputs provided a close representation of the pseudo-calibration observations.

## Habitat Analysis

The Middle Oconee River is very biodiverse with over 27 species of fish caught at the study site (personal communication, Mary Freeman, US Geological Survey). For comparison, the entire Colorado River Basin has less than 30 species of native fish. The biodiversity of southeastern streams makes a species-by-species habitat assessment prohibitively difficult and impractical (Bowen *et al.*, 1998).

To fill this gap, Freeman et al. (1997) developed generalized habitat suitability criteria for nine fish species by including depth, velocity, substrata type and cover. This study addressed three key types of habitat with accompanying criteria for depth and velocity: shallow-fast, deep-fast, and shallow-slow (Table 4.1).

Using these criteria, habitat suitability was predicted over 79 values of river discharges: $10-200$ cfs (by 10 cfs), $200-1,000$ cfs (by 20 cfs), and $1,000-20,000$ cfs (by $1,000$ cfs). For each discharge, HEC-RAS was executed under steady-state conditions, and spatially explicit velocity and depth distributions were generated. A Python script was then applied in ArcGIS® (version 10.2.2) to calculate wetted usable area (i.e., total available habitat) and suitability for each habitat type for each of the 79 discharge simulations (Example shown in Figure 4.6).

Each of the 304 environmental flow regimes were then assessed on the basis of habitat. First, each of the four habitat types was assessed at the average annual discharge for each flow regime. This provides a general perspective on the rate of decline in available habitat with increasing water withdrawal. Second, magnitude-frequency analysis was used to estimate the amount of habitat across the entire range of the flow regime (Figure 4.7). For each environmental flow alternative, a frequency distribution of all discharge values was obtained. A habitat rating curve was developed from the HEC-RAS simulations. The product of the amount of habitat (i.e., magnitude) and probability of occurrence (i.e., frequency) provides a relative measure of the time-weighted amount of habitat (i.e., the effectiveness). The area under this curve is the total amount of habitat provided by the entire flow regime over the 60-year simulation period, which provides a more holistic measure of habitat than habitat quantity at the average annual discharge.

## 4.3 Results

Simulations of environmental flow alternatives resulted in different types of hydrologic alteration for each of the three families of methods (AMF, MMF, POF). Impacts to the time series of flow were noticeably different for the minimum flow strategies and percentage-based strategy, even at similar levels of total withdrawal. The minimum flow approaches led to periods of âĂIJflat-lining,âĂİ whereas the POF approach maintained variability throughout the simulation (Figure 4.8, top). In addition to changes in river hydrographs, the strategies led to different withdrawal volumes in each of the 60-years simulated (Figure 4.8, bottom).

Hydraulic and habitat simulations provided a mechanism to construct habitat rating curves for each of the four types of habitat assessed here (total, shallow-fast, deep-fast, and shallow-slow; Figure 4.9). As expected, total habitat increases with increasing discharge. However, the distribution of habitat types changes dramatically over the range of discharges simulated. In particular, shallow-fast habitat appears only during a narrow band of discharges and represents a small portion of the total habitat, a notable outcome given the disproportionate biodiversity of these "shoal" conditions (Travnichek & Maceina, 1994). The mosaic of habitats is, thus, differentially affected with changes in the environmental flow regime.

These simulations provide a mechanism to assess trade-offs between municipal water supply and habitat provision under the three environmental flow schemes (AMF, MMF, POF). For instance, for any given level of water withdrawal, which alternative provides the most habitat, and for any given level of habitat, which alternative provides the most water? While the average municipal water yield across the scenarios is of interest to many applications, municipal water suppliers must often focus on the outcome with the greatest potential for societal impact, the minimum water yield across the 60-year simulation. This represents the "worst case scenario" for water supply outcomes, and thus, is used in assessing trade-offs with habitat (Figure 4.10). The three environmental flow alternatives are compared on an equal withdrawal basis in an effort to find the most efficient alternative. From the perspective of total available habitat and deep-fast habitat, both the average discharge output and the frequency-weighted output indicate the percent-of-flow approach to be the most efficient. However, results are mixed for the shallow-fast and shallow-slow habitats with all three environmental flow alternatives indicating efficiencies at different withdrawal rates.

Key differences emerge in the findings based on average annual discharge (Figure 4.10, top) or a magnitude-frequency analysis (Figure 4.10, bottom). First, total habitat is consistently over-predicted by the average discharge method. This is an expected outcome given that flow frequency distributions are often highly skewed, which leads to a mean discharge much greater than the median discharge (e.g., 521 cfs vs. 350 cfs for the Middle Oconee River). This skewed distribution is accounted for when incorporating the frequency of flows via effectiveness analysis, while average discharges can indicate a false sense of the quantity of habitat available. Second, only tracking average discharge can mask nuanced effects associated with alternative environmental flow regimes. For instance, the shallow-fast habitat assessments with magnitude-frequency analysis show a non-monotonic response, potentially due to changes in low flows as well as central tendencies. Third, the relative ranking of environmental flow alternatives shifts depending on whether average or frequency-weighted conditions are used.

## 4.4 Discussion

This study addressed two primary objectives. Our first major objective was to demonstrate a suite of methods for comparing simple environmental flow alternatives and developing flow thresholds, which apply a common approach to habitat analyses. We coupled three existing families of methods to accomplish this objective. Hydrologic simulation is a common tool for examining potential consequences of operational changes at water infrastructure (Klipsch & Hurst, 2013, McKay, 2015) and watershed-scale planning (LaFontaine *et al.*, 2015). Habitat analysis has a deep history in water resource management and environmental flow analysis, which includes a suite of techniques such as the Physical Habitat Simulation (PHABSIM), Instream Flow Incremental Methodology (IFIM), the Riverine Community Habitat Assessment and Restoration Concept (RCHARC), the Ecosystem Functions Model (HEC-EFM) (Hickey & Fields, 2009), and a variety of other methods (Reviewed in McKay 2013). Magnitude-frequency analysis (Wolman & Miller, 1960, Doyle *et al.*, 2005) then provided a mechanism to synthesize outputs from the hydrologic simulations and the habitat models. We coupled simple forms of these techniques using hydrologic simulation (McKay, 2015, McKay *et al.*, 2016), generalized habitat criteria (Freeman *et al.*, 1997), and a basic form of effectiveness analysis (Doyle *et al.*, 2005) to create a useful analytical framework easily adaptable to other study sites. Currently, standardized USACE platforms can perform many of these functions (e.g., HEC-ResSim, Klipsch and Hurst 2013; HEC-EFM, Hickey and Fields 2009).

Our second major objective was to examine the role of variability in setting flow management thresholds by comparing environmental flow recommendations developed using "typical" river levels (i.e., long-term averages) with those developed using a frequency-weighted approach (i.e., effectiveness analysis). Environmental flow and river restoration analyses often use average discharge conditions when assessing the benefits of restoration actions. However, this analysis shows that average conditions are not necessarily indicative of the effects of environmental management actions. The effectiveness analysis approach applied here addresses some (not all) of these deficiencies by incorporating the range of variability in discharge along with the event frequency. However, timing, duration, and rate-of-change of flows can also be crucial to ecological functions (Poff *et al.*, 1997) and were ignored in this analysis. We also assumed the bathymetry of the stream to be unchanging.

## 4.5 Conclusion

This study contributes to a growing body of information about the effects of river flow regimes on the Middle Oconee River ecosystem (Nelson & Scott, 1962, Grubaugh & Wallace, 1995, Katz & Freeman, 2015, McKay, 2015, McKay *et al.*, 2016). A number of these studies have focused on the consequences of different environmental flow alternatives to, for example, monthly or annual $7Q10$ flows. This body of evidence is indicating that overall, percent-of-flow (POF) approaches appear to have fewer ecological impacts than minimum flow approaches. However, results shift depending on the process investigated, and in some scenarios minimum flow approaches are functionally equivalent to percent-of-flow approaches (McKay, 2015, McKay *et al.*, 2016). This research suggests that data-driven methods such as frequency-based approach is beneficial in providing insights in search of alternatives to $7Q10$ flow regulations to improve water management sustainability in the region. Although some outcomes are highly sensitive to changes in low flow conditions, many of the species in the river exhibit remarkable resilience to drought and low flows (Katz & Freeman, 2015). Additional research will help identify which environmental flow regimes are most efficient across ecological outcomes (e.g., fish communities, primary production, etc.).

In order to take into account the impact of development on the environment, it is essential to study flow regimes and trade-offs involved in water management. Here, a new coupling of analytical techniques is presented, which helps incorporate natural variability into environmental flow studies. In doing so, we demonstrate the importance of hydrologic variability, not only relative to ecological outcomes, but relative to water management decision making.


## Acknowledgements

Research presented in this technical work was developed under the Ecosystem Management and Restoration Research Program (EMRRP) and Georgia Water Resources Institute (GWRI) Program. The USACE Proponent for the EMRRP Program is Mindy Simmons and the Technical Director is Al Cofrancesco. Bruce Pruitt and Lucas Montouchet collected 2013 survey data, and Tom Prebyl assisted with geospatial analyses. Mary Freeman (US Geological Survey), Caitlin Conn, Seth Wenger (University of Georgia), John Hickey (USACE Hydrologic Engineering Center), Sarah Miller (ERDC-EL), Mick Porter (USACE Albuquerque District), and Brian Zettle (USACE Mobile District) graciously reviewed this document.

Table 4.1: Habitat suitability criteria and representative taxa observed in the Middle Oconee River.

| Key Habitat | River Depth | Flow Velocity | Representative Taxa |
|---|---|---|---|
| Shallow-Fast | $\leq$ 35 cm | $\geq$ 55 cm/s | *Nocomis leptocephalus* (bluehead chub) *Notropis hudsonius* (spottail shiner) |
| Deep-Fast | $\geq$ 35 cm | $>$ 45 cm/s | *Micropterus Salmoides* (largemouth bass) |
| Shallow-Slow | $<$ 35 cm | $<$ 35 cm/s | *Lepomis* (bluegill and sunfish) |

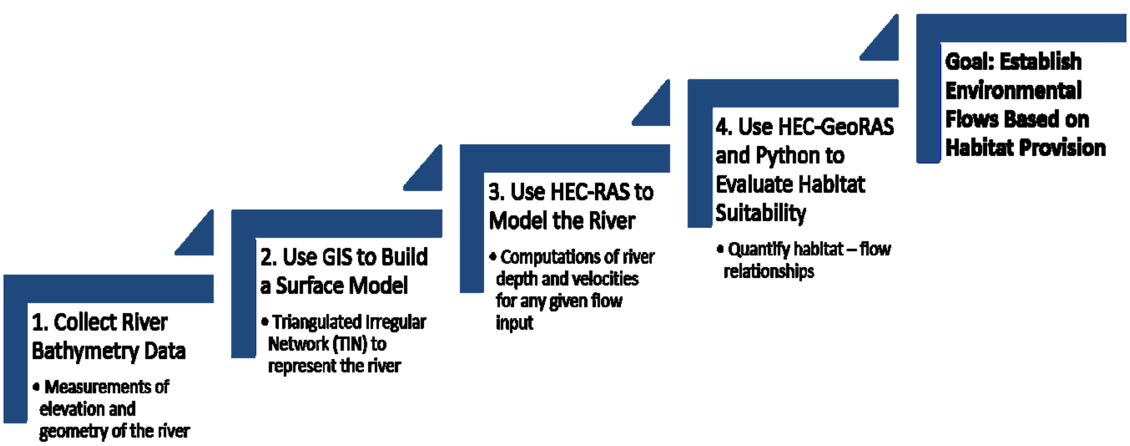



Figure 4.1: Research Framework for Quantifying Habitat-Flow Relationship.

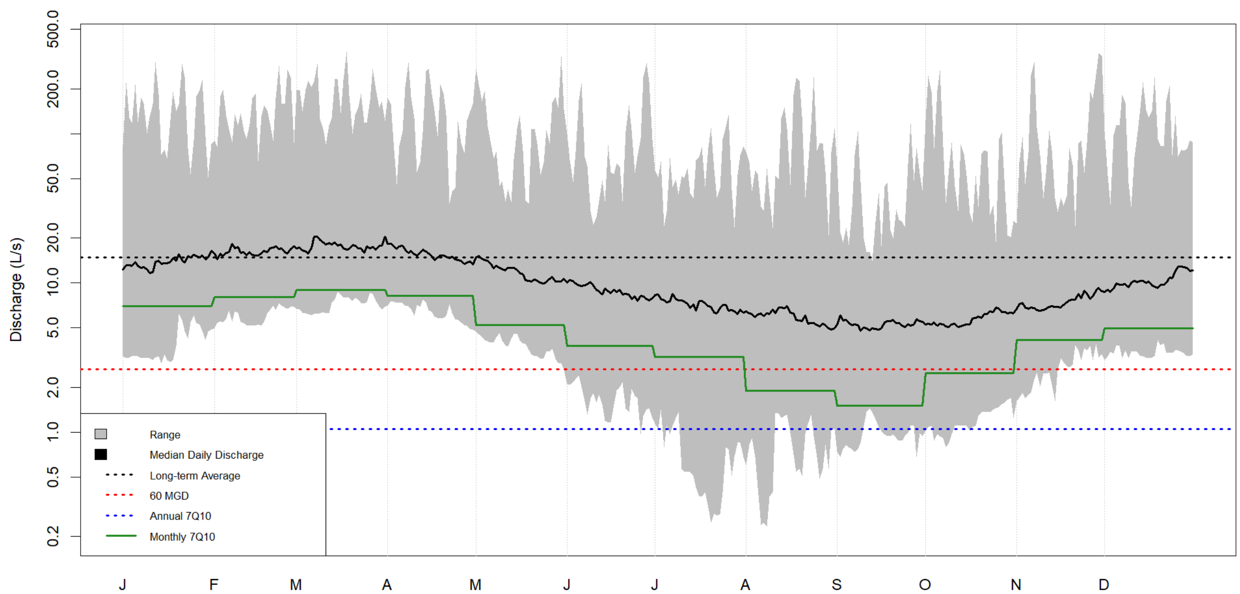



Figure 4.2: Long-term, minimally altered hydrograph on the Middle Oconee River near Athens (1938-1997). The shaded area represents the lowest and highest discharge observed on each day of the year, the solid black line is the daily median, the dashed black line is the long-term mean, the red dashed line is the pump capacity for Bear Creek Reservoir, the blue dashed line is the Annual $7Q10$ level, and the solid green line is the monthly $7Q10$.

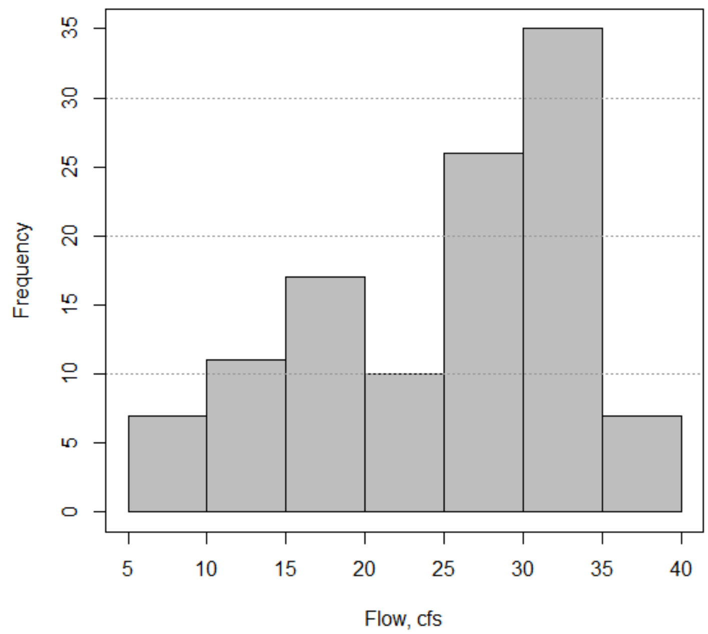



Figure 4.3: Histogram of Low Flows below the Annual 7$Q$10 level on the Middle Oconee River near Athens during 1938-1997 (prior a construction of the Bear Creek Reservoir).

Figure 4.4: Bathymetry data collection on the Middle Oconee River near Athens, Georgia.

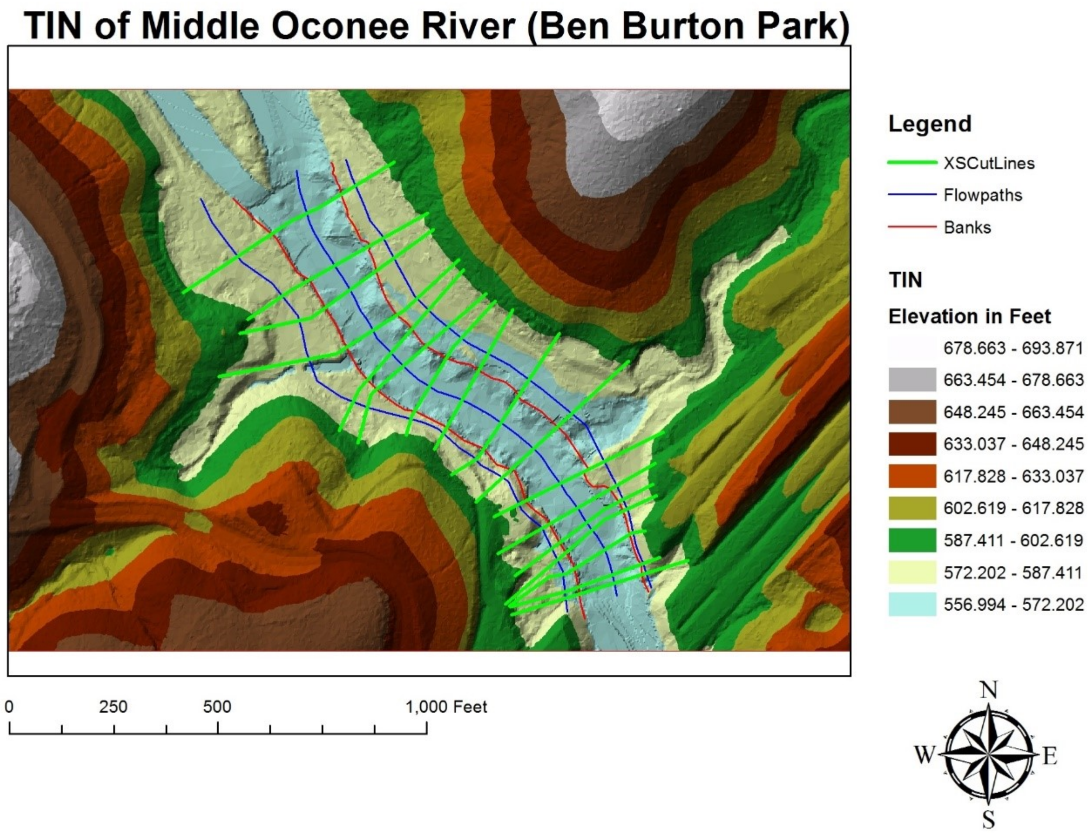



Figure 4.5: Map of the reach on the Middle Oconee River near Athens, Georgia. Surveyed cross-sections are shown in green, general flow paths of the main channel and floodplains as blue lines, and bank demarcation points as red lines.

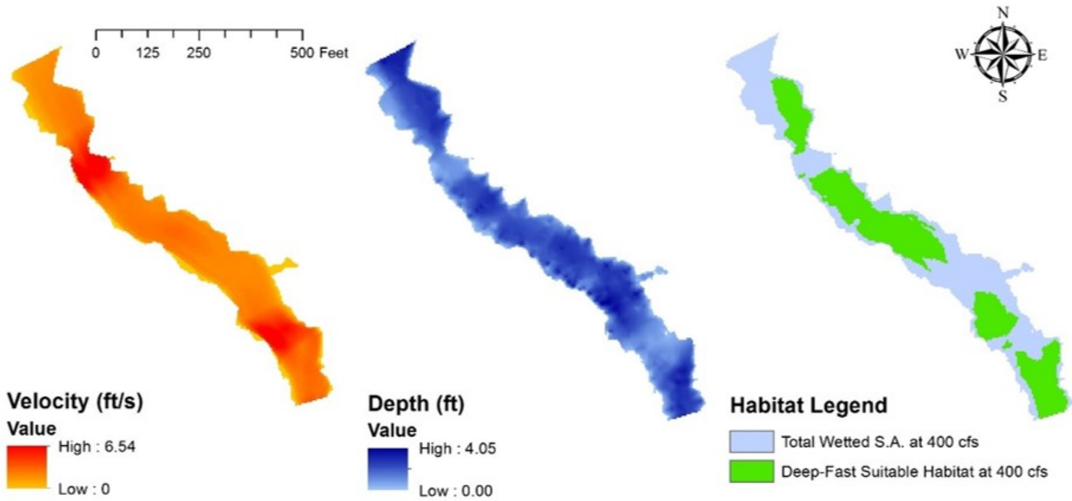



Figure 4.6: Example of spatially explicit outputs for hydraulic and habitat models at 400 cfs: (A) velocity, (B) depth, and (C) deep-fast habitat suitability.

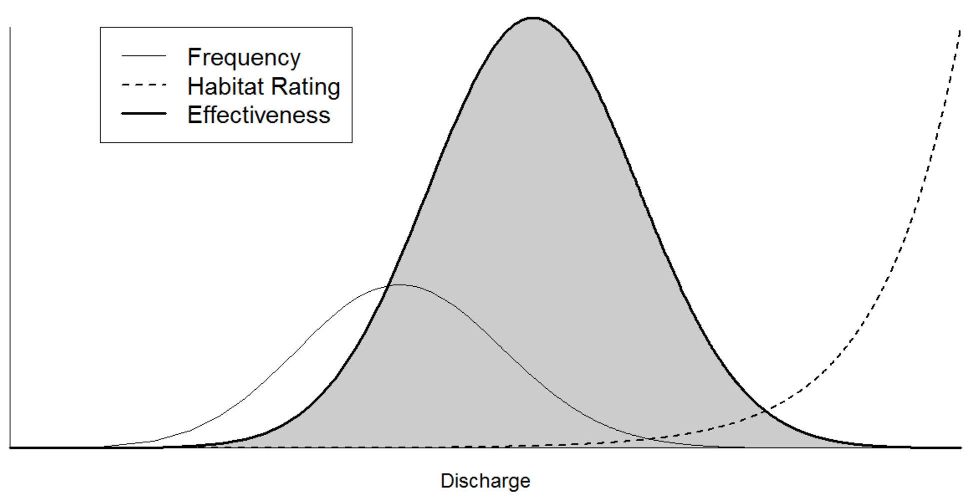



Figure 4.7: Conceptual depiction of effectiveness analysis (i.e., magnitude-frequency analysis).

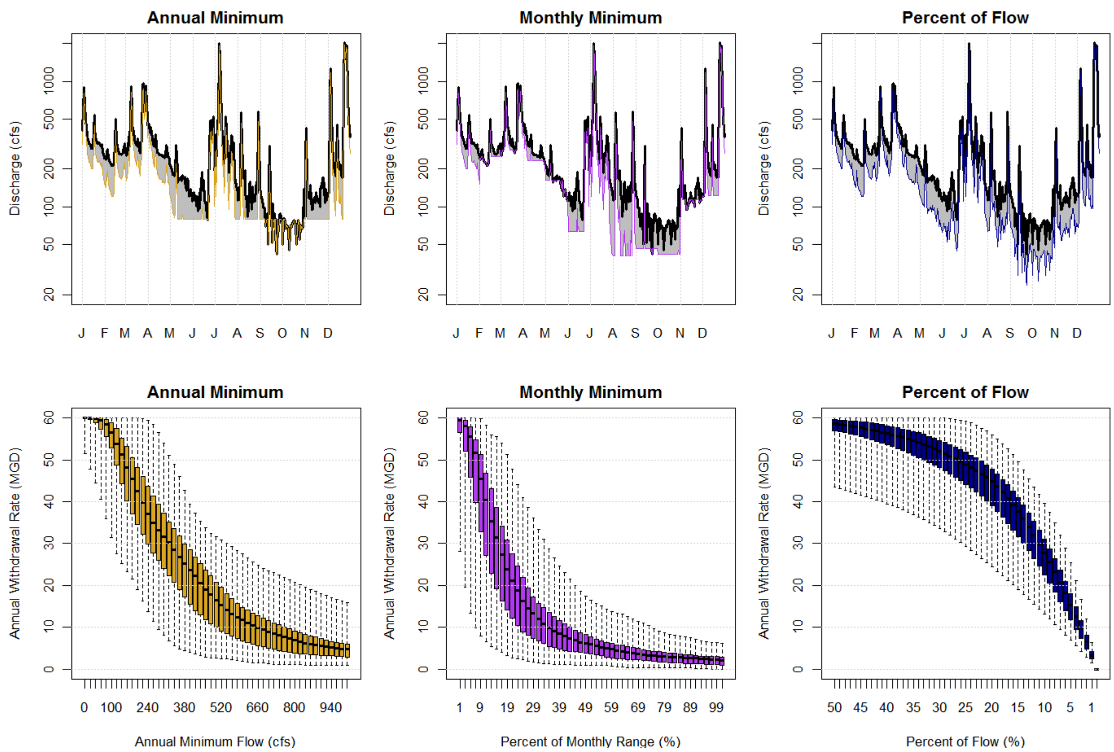



Figure 4.8: Effects of environmental flow alternatives. (top) Example of hydrographic effects for the sample year 1941. All flow management alternatives provide similar levels of average water withdrawal, i.e., 55.5, 55.4, and 55.5 MGD, respectively. (bottom) Variability in municipal water withdrawal over a range of environmental flow thresholds.

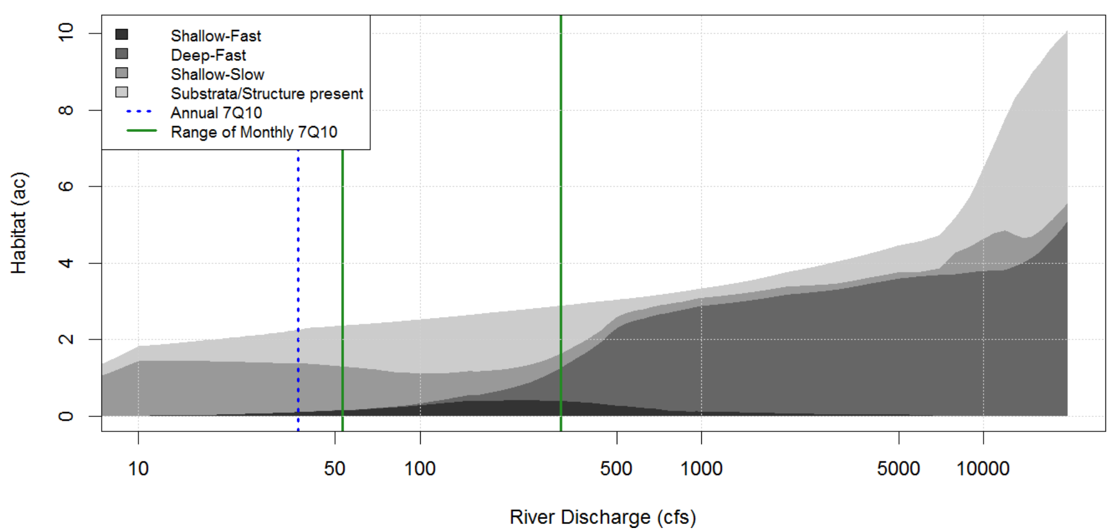



Figure 4.9: Cumulative habitat rating curves over the range of discharges observed.

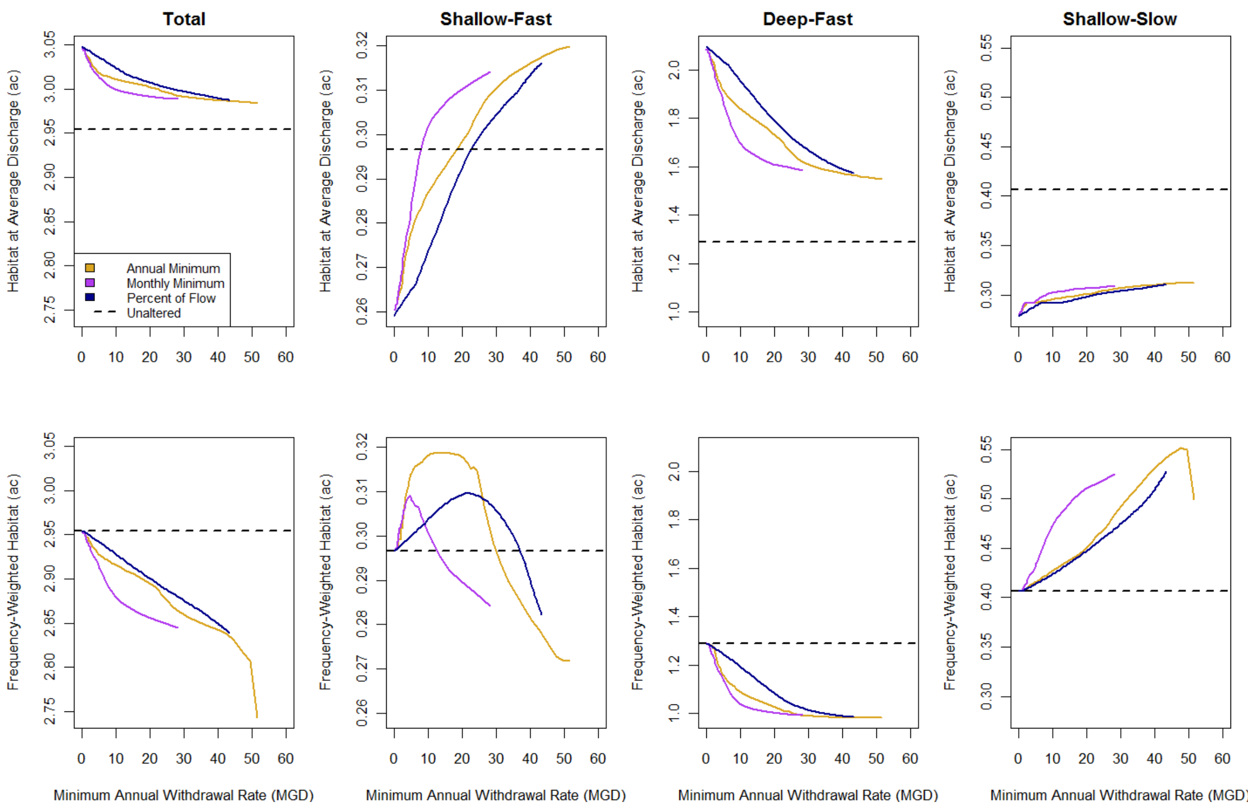



Figure 4.10: Comparison of environmental flow alternatives across total habitat and three distinct habitat types. (top) Habitat computed only at average discharge. (bottom) Habitat computed as a frequency-weighted quantity using effectiveness analysis.

# Chapter 5

# Applying environmental flow analysis and time series analysis for water management and habitat analysis along selected reaches[1]

[1]N. V. Bhattacharjee, J. R. Willis, S. K. McKay and E. W. Tollner. To be submitted to *Journal of Environmental Management*

# Abstract

According to the Middle Ocmulgee Regional Water Plan, the population in the Middle Georgia region is projected to double by 2050 which will result in a water withdrawals increase of 38% and wastewater return flows increase of 62%. Therefore, it is important to examine trade-offs in water management to support both sustainable ecosystems and economic development of the region.

In this study, we focused on two reaches: the first river stretch is upper reach near the Ocmulgee National Monument at Macon; the second part is lower reach near Bullard's Landing at Warner Robins. We studied the influence of power plant operation on the study sites and associated fish habitat. We first conducted a hydrographic field survey to obtain bathymetry data using kayak-mounted sonar-based GPS mapping system. We used the Hydrologic Engineering Center's River Analysis System (HEC-RAS) developed by the US Army Corps of Engineers. The analysis on how reservoir operation can influence local hydrology and fish communities helped quantifying economic and ecological trade-offs involved in water management. We applied the methods and techniques from our previous study on the Middle Oconee River (Bhattacharjee *et al.*, 2017) to a larger scale with a goal of bringing insights and new knowledge in improvement of water management on the Ocmulgee river. This study demonstrates the importance of hydrologic variability for environmental flows study and ability of effectiveness analysis to be adapted in other study sites.

In support of the Ocmulgee Water Trail Initiative, we evaluated existing landings and worked on specific recommendations for new landings that can improve recreational access and tourism between Macon and Hawkinsville. Thus, we identified high potential locations for future improvements which will guide the Ocmulgee National Park Initiative to take further strides towards establishing the Ocmulgee Mounds National Historical Park.

## 5.1 Introduction

A wide range of human activities has led to rivers being deemed the earth's most damaged ecosystem, losing species at a greater rate than terrain and marine ecosystems (Dudgeon *et al.*, 2006). It is imperative that rivers are managed in a more natural, sustainable way that balances the needs of both the aquatic ecosystems and the human livelihoods that rely on them. Environmental flow regimes are offered as at least a partial solution to some of the freshwater challenges that are currently in our midst. According to the Middle Ocmulgee Regional Water Plan, the population in the region is projected to double by 2050 which will involve water withdrawals increase by 38% and wastewater return flows by 62% by 2050 (RWP, 2011). Therefore, it is important to examine trade-offs in water management to support a sustainable ecosystem of the region.

A variety of studies have been conducted on riverine systems to evaluate the impacts that resource management and other human factors have had on aquatic ecosystems. In Georgia, a trade-off study for water management strategies along the Middle Oconee River has been conducted by McKay, 2015. In his study, McKay examined the quantity of aquatic habitat and water withdrawals associated with several environmental flow types and constructed a decision making framework for the Middle Oconee River near Athens, Georgia (McKay *et al.*, 2016). Furthermore, Gibson et al. (2005) conducted a study of two river basins (Cle Elum River, Washington and Chattahoochee-Apalachicola River Basin, Georgia and Florida) to investigate the impact of future climate scenarios on river ecosystem. They demonstrated significant changes in flow regimes and aquatic habitat under various climate scenarios (Gibson *et al.*, 2005). Additionally, a comparison study of two rivers in Northern Michigan (Carp Lake River and Little Black River) was conducted to examine ecological effects of agricultural development on stream habitat and nutrient input (Dillon, 2013). In Alabama, Swinson (2014) studied the Tulotoma snail habitat along the Coosa River. The main focus of his study was contour generation using HEC-RAS software based on geo-referenced bathymetry of the Coosa River (Swinson, 2014). Moreover, Yao and Georgakakos introduced a concept of adaptive water resources management in their study of Folsom Lake, California. The adaptive system demonstrated reliable forecasts for better reservoir performance when compared to traditional ones (Georgakakos *et al.*, 2012, Yao & Georgakakos, 2001). Also, Chen et al. (2015) addressed trade-offs for better water management of the Jordan River in the Middle East and the Colorado River in the western United States. They addressed similarities and differences of the two river basins by considering various factors, such as increasing water demand and supply and environmental flow demand (Chen *et al.*, 2015).

Previous research shows a relationship between discharge and fish biomass based on a study of Austrian rivers (Moog, 1993). Freeman et al. (1997) developed habitat suitability criteria for nine fish species by including depth, velocity, substrata type and cover (Freeman *et al.*, 1997). For Eastern warm water US rivers, researchers suggest that it is not practical to collect habitat data due to a high number of the species present (Bowen *et al.*, 1998). Therefore, researchers used more generalized criteria for habitat analysis rather than species-specific criteria (Bain, 1995).

The Ocmulgee River harbors high aquatic biodiversity, and protecting these species is of high priority for the region. The river is well-known as a home for largemouth bass, bluegill (*coppernose bream*), redear sunfish (or *shellcracker*) and flathead catfish (or *appaloosa cats*). Additionally, other species such as redbreast sunfish, black crappie and chain pickerel represent local favorites (known as game/sport fish species). In this study, we applied the concepts and techniques on habitat analysis described in Chapter 4 (Bhattacharjee *et al.*, 2017). One of the important fish species is Robust Redhorse species. They have been stocked in the Ocmulgee river and they almost certainly occurred there historically. The other big river species are Shortnose and Atlantic sturgeons and American Shad (based on personal communication with Christopher Skelton, fisheries biologist).

We examine trade-offs in water management between ecological impacts and economic development by looking at the historical streamflow discharge data which are then examined relative to local hydraulic conditions and aquatic habitat needs. In order to model hydraulics of the river, we use the Hydrologic Engineering Center's River Analysis System (HEC-RAS) developed by the US

Army Corps of Engineers (Brunner, 2001).

The Ocmulgee Water Trail encompasses approximately 320 km of water trail stretching from Macon to the river's confluence with the Altamaha river. The Ocmulgee River Water Trail Partnership consists of eleven counties stretching from Macon to Lumber City and including Pulaski County and the town of Hawkinsville. The National Park Service found that paddlers spent between $27 – $63 per day in communities along water trails. Middle Georgia Regional Commision seeks opportunities to improve connectivity in the Ocmulgee River corridor (Commision, 2012). In this study, we also focus on identifying potential future landings to maximize the connectivity of the trail by considering reach from Macon to Hawkinsville.

At the time of bathymetry data collection in middle May 2016, we spotted several alligators along the lower reach. Local fisherman and raft guides told us that alligator sightings are rare in the upstream reach but quite common in the lower reach. Therefore, it was only feasible to collect bathymetry data from Amerson Park to the Ocmulgee National Monument (from Landing 1 to 2); and from Bullard's Landing to Knowels landing (Landing 3 to 4). Bathymetry data collection from Landing 2 to Landing 3 was not possible as the distance between these two access points was 50 km (therefore, we proposed potential location of future landings in Section 5.3, Figure 5.6). Data collection below Landing 4 was not advisable due to presence of alligators. The detailed description of the bathymetry data collection is given in Section 5.2.

The main objectives of our study are to:

1. examine flow regimes relative to local hydraulic conditions and aquatic habitat needs using hydrologic and hydraulic models;

2. identify potential landings to support the Ocmulgee Water Trail Initiative using geospatial analysis.

## 5.2 Methods

The USGS gage number closest to our study area is 02213000 near Macon, GA. River flow is currently affected by Lloyd Shoals Dam (built in 1910) that is located near Jackson Lake. The long-term data of stream flow is available from 1894 till present with missing data from August 1912 to September 1929. In this work, we considered flow from during $1894 - 1909$ as an unaltered flow (16 years of long-term time series data). It is permitted to withdraw a maximum of 110 million gallons per day (4.82 L/s or 170.21 cfs; Georgia EDP Permit Number $011 - 0590 - 02$); this level is marked as a red dashed line in Figure 5.1. Summary of descriptive statistics of an unaltered flow is given in Table 5.1.

Figure 5.1 provides a time series of an unaltered flow along with daily median, long-term mean, lowest and highest discharge values as well as $7Q10$ flow. This hydrograph for the Ocmulgee river suggests higher flow variability when compared to the hydrograph for the Middle Oconee river (Figure 4.2 in Chapter 4 (Bhattacharjee *et al.*, 2017)). The driest month, with the minimum monthly flow of 8.5 L/s or 300.97 cfs was October. This record is below annual $7Q10$ level of 11.6 L/s or 410 cfs (Figure 5.1). There were a total of 73 days when the discharge was lower than the annual $7Q10$ level during the 16-year study period ($1894 - 1909$). The low flow events were recorded during September - October months (Figure 5.1).

### Geospatial Analysis

The Ocmulgee Water Trail provides opportunities for various recreational activities, such as wildlife observation, fishing and canoeing. It is anticipated to be a premier destination for paddlers and river enthusiasts (http://ocmulgeewatertrail.com/). In support of the Ocmulgee Water Trail Initiative, we evaluated existing landings and worked on specific recommendations for new landings that can improve recreational access and tourism between Macon and Hawkinsville. We performed geospatial analysis in ArcGIS® (version 10.2.2) and used the site evaluation criteria given in

Table 5.2.

The above criteria were considered based on interviews with groups from the Ocmulgee Water Trail Partnership, faculty from Carl Vinson Institute of Government, representatives of the Chamber of Commerce in Macon, Middle Georgia Regional Commission and a previous report for UGA-Archway Partnership delivered by Eason and Hameduddin, 2013. Based on data of the existing landings, we propose potential locations for future landings to improve connectivity between landings 2 and 3; 6 and 7. Therefore, we planned to collect bathymetry data from landing 1 to landing 2 and from landing 3 to landing 6. However, we were not able to use the survey equipment from landing 4 to landing 6 due to presence of alligators (as we mentioned in Section 5.1).

## Bathymetry data collection

The bathymetry data collection process relied on the use of a kayak "pyranha" and the coupling of $TrimbleGeo7X$ GPS and sonar technologies. A detailed description of the sonar system can be found in Swinson, 2012. A Real-Time Kinematic (RTK) satellite navigation system was secured near the stern of the kayak using a rod (Figure 5.2). The RTK device received real-time GPS corrections from a reference station that resulted in sub-meter horizontal and vertical accuracies. The satellite system provided geospatial (X,Y) coordinates along with water-surface elevations for each coordinate. A sonar depth-finder system was used to measure depths from the water surface to the river bottom. A multiplexer device received and stored the inputs from the RTK and sonar systems. Matched via timestamps and used in conjunction, the RTK and sonar measurements combined to provide accurate estimates of the river bottom elevation.

River bathymetry data were collected as the kayak was maneuvered in a downstream direction. At every 150 − 200 feet, the kayak was directed from one bank to the other in as straight of a line as possible that was perpendicular to the longitudinal direction of flow. After collecting the bathymetric data with the kayak-mounted satellite and sonar system, RTK data consisted of points with a precision of less or equal to 0.3 meters. The lower precision points occurred when the GPS system locked onto a low number of satellites and/or communication with the reference system was lost. This additional data filtering step served to increase the reliability of the measured river bottom elevations.

During the bathymetric survey, there were difficulties encountered while collecting data. Banks with vegetation and biomass protruding outwards into the channel, particularly in the lower reach, sometimes obstructed the kayak from collecting data near the banks. This vegetation also disrupted the satellite signals at times. Furthermore, there were a few stretches of the river that were partially blocked with fallen trees and snags. The combination of debris and river currents in these areas were difficult to maneuver around with a kayak.

Noteworthy differences were observed in the Ocmulgee River between the upper reach stretching from Amerson Park to the Ocmulgee National Monument and the lower reach stretching from BullardŠs Landing to Georgia State Route 96. The upper reach was generally wider and shallower and consisted of a rockier substrate and clearer water. The river portion adjacent to the Ocmulgee Monument was deeper and more braided, with several beach areas with sediment deposition. While collecting river data at the lower reach, the observed characteristics of the river were strikingly different. The river channel was considerably deeper and narrower with a softer bottom and more vegetated banks. The turbidity of the water was higher and was brown in color. In addition, the reach contained a noticeably more obstructions in the main channel such as fallen trees and snags. It is also interesting to note that several alligators were spotted along the lower reach whereas none were spotted upstream near Macon. After speaking with local fisherman and raft guides, a consensus was reached that alligator sightings are rare in the upstream reach but quite common in the lower reach even though the two reaches are separated by less than 20 miles.

## Hydraulic Modeling

Elevation data pertaining to the banks and floodplains of our study region were obtained from a digital elevation model (DEM) derived from TerraASTER satellite data (https://earthdata.nasa.gov/).

The collected bathymetric data in conjunction with the acquired DEM data of the bank and flood-plains were used to create a triangulated irregular network (TIN) that closely resembles the main channel, bank slopes, and flood-plain surface of our reach. Figure 5.3 represents a model-builder of the data processing steps performed in ArcGIS® (version 10.2.2). Geoprocessing of a long reach required splitting the river stretch to avoid run time errors when using HECGeoRAS tool for generating depth and velocity profile maps. Therefore, for hydraulic modeling and habitat analysis, we considered length of each reach of approximately two miles. The Ocmulgee River reach nearest to the Ocmulgee National Monument is located near Macon, within Bibb County; wherein the Ocmulgee River reach close to Robins Air Force base is located near Warner Robin within Houston county.

Prior to importing the TIN surface into HEC-RAS for hydraulic analysis, we used HEC-GeoRAS (version 10.1), an ArcGIS extension toolbar to digitize key components such as the river centerline, bank locations, flow paths and cross-sections. Following standard convention, flow paths in the flood plains were estimated to be at the center of mass flow between the top of the bank and the extent of the floodplain (roughly one-third of the distance from the banks and two-thirds from the floodplain extent) as shown in Figure 5.4. The cross-sections were bent or “doglegged" to intersect the left overbank flow path, main channel flow path, and right overbank flow path perpendicularly. The Federal Emergency Management Agency (FEMA) 100 year flood map was used to ensure that the cross-sections were digitized in a manner that spanned the entire extent of the floodplain. The final TIN surface and digitized features representative of the Ocmulgee River near the National Monument are given in Figure 5.4. Figure 5.5 represents TIN surface and digitized features of the Ocmulgee River near Warner Robins.

To model hydraulics of the river, including the velocities and depths associated with a given flow rate, we used the Hydrologic Engineering Centers River Analysis System (HEC-RAS) software. Once opening the HEC-RAS import file (created from HEC-GeoRAS) in HEC-RAS, a series of adjustments were made in order to calibrate the model as accurately as possible. We calibrated the model based on the following information. While gathering the bathymetric survey data, we measured the elevation of the water surface at each bank of each cross-section along with the date and time. We then retrieved discharge data from a nearby USGS gage station measured at the same day/time of data collection. Thus, with doing a flow analysis at a specified flowrate in HEC-RAS, we knew that the water level in the model should closely reflect what we measured it to be in the field. For the flow analysis, steady flow and subcritical flow were assumed. Thus, the major inputs that were adjusted to tune our model in HEC-RAS were Manning's 'n' roughness coefficients and the downstream riverbed slope. The choice of ManningŠs *n* coefficients were guided from standard tabulated values found in the text ŞOpen Channel HydraulicsŤ (Chow, 1959). ManningŠs *n* values used for all of the cross-sectional floodplains and main channel sections fall within accepted ranges.

The deepest measured river points (thalweg) within each cross-section were used to determine a composite slope for the entire reach. Starting with the most upstream cross-section, the cumulative distance from the first thalweg point to subsequent thalweg points downstream was calculated using the latitude and longitude coordinates of the points. The cumulative distance values were plotted against the measured elevations of each of the thalweg points and a line of best fit was employed.

## Flow Regimes and Habitat Analysis

We consider four scenarios of municipal water withdrawal and environmental flow requirements: Unaltered, Annual Minimum Flow (AMF), Monthly Minimum Flow (MMF) and Percent of Flow (POF) described in detail in Chapter 4 (Bhattacharjee *et al.*, 2017). We modify the unaltered hydrograph for the period of $1894 - 1909$. Flow was abstracted at a maximum withdrawal rate of 110 MGD in accordance with DNR policy. Result of each simulation was a 16-year record of daily discharge and daily water withdrawal time series used in habitat trade-off analyses.

We used a generalized criteria for habitat analysis rather than species-specific criteria (Bowen *et al.*, 1998, Freeman *et al.*, 1997). Similarly to the study on the Middle Oconee river, we analyzed different flow regimes under the following settings: $100 - 4,000$ cfs (with intervals of 50 cfs), $4,000 - 10,000$

cfs (with intervals of 500 cfs), $10,000 - 100,000 cfs$ (with intervals of $5,000$ cfs). As a result, we calculated wetted usable area for each fish assemblage based on the outputs from HEC-RAS model based on 109 input values of river discharge. For each input value, we obtained velocity, depth and habitat maps using a Python script developed and presented in Appendix D.

### Spectral Analysis in Time Series

The main idea behind spectral analysis is to decompose a time series into a combination of sinusoids ($sin$ and $cos$). It is referred as an analysis in the "frequency" domain. Periodogram is a graph that provides information about the periodic components of the time series. It shows the relative strengths of the various frequencies for explaining the variation in the time series. Periodogram is used to identify the dominant periods $T$ (or frequencies $f = \frac{1}{T}$) of the time series (Haan, 2002). For example, a seasonal period (90 days) would have frequency of $0.011 = \frac{1}{90}$, monthly period would have frequency of $0.033 = \frac{1}{30}$, bi-weekly period would correspond to frequency of $0.071 = \frac{1}{14}$ etc.

Periodogram is a rough estimate of a population spectral density and it is very noisy. Therefore, we used two smoothing methods for estimating the spectral density: nonparametric estimation (using a centered moving average procedure) and parametric estimation (using autoregressive (AR) model). We used *spec.pgram* and *spec.ar* functions available in the R statistical software (version 3.2.5) in order to estimate the spectral density using the nonparametric and parametric estimations, respectively.

## 5.3 Results and Discussion

### Geospatial Analysis

We examined the site evaluation criteria described in section 5.2 for the purpose of assisting the Ocmulgee Water Trail Partnership in identifying the best sites for future growth and expanding tourist activities. As a result, we identified four high potential locations for future improvements (marked as black squares in Figure 5.6). Local community has been given database with ownership data, geographic data and potential take-off landings location data. This information will allow them to validate and select the potential location by working with land owners based on provided parcel data.

## Hydraulic Modeling and Habitat Analysis

We attempted to generate depth and velocity profile maps for a longer reach than the stretch we analysed in our previous work (Bhattacharjee *et al.*, 2017). Therefore, in order to avoid run errors while using HECGeoRAS tool, we considered the length of study site of approximately $2-3$ miles. For all 109 flow regimes, it required more than 50 hours to generate depth and velocity profiles in HECGeoRAS and process them in Python for each river stretch prior conducting habitat analysis.

Figure 5.7 illustrates the change of habitat area near Macon with respect to the range of flow discharges measured within the 16 year analysis period. Shallow-fast habitat takes a very small portion of the wetted area, followed by shallow-slow habitat. Area for deep-fast and substrata/structure dependent habitat types increases as the flow discharge increases. However, the results suggest the dominance of Substrata/Structure dependent habitat category over shallow-fast, deep-fast, shallow-slow categories. The habitat rating curves demonstrate the importance of information on substrata, structure in the river (Table 2.1) that will allow introducing another key habitat in the classification of fish assemblages. The distribution of different types of habitat changes significantly especially when flow is above 2000 cfs.

In order to look at low flow discharges, we zoom Figure 5.7 to look at the habitat distribution near Macon over the range of 250 to 2000 cfs. Shallow-fast habitat occupies very small portion of the total habitat and it is very low for the flows above 1000 cfs. Shallow-slow habitat almost stays constant accross the $250-2000$ cfs range wherein deep-fast and sustrata/structure habitat increases as discharge increases (Figure 5.8).

We compare the habitat distribution near Macon with the habitat distribution near Warner Robins (Figure 5.9). Overall, the trends look similar over the range of 250 to 2000 cfs. However, it is noticeable that shallow-fast habitat near Warner Robins occupies very small part of the wetted area and it is much smaller when compare to shallow-fast habitat near Macon. Additionally, the rate of increase in each habitat (especially, substrata/structure dependent habitat type) near Warner Robins is larger than that near the Ocmulgee National Monument area.

Figure 5.10 includes results from effectiveness analysis when we computed habitat as a frequency-weighted quantity. The graph shows similar trends for total and deep-fast habitat. However, both shallow-fast and shallow-slow suggest opposite response to withdrawal rate increase compared to the results from average discharge analysis. Shallow-fast type appears to increase as withdrawal rate increases where either MMF or AMF provides larger habitat area when compared to POF. Shallow-slow habitat area exponentially decreases until it becomes almost a flat line (at 20 cfs for AMF and MMF; 60 MGD in case of POF). Across all habitat types, POF seemed to provide more habitat, except for the shallow-fast category (Figure 5.10).

Habitat calculated only at average discharge for AMF, MMF and POF alternative regimes supports results in our previous study (Figure 4.10, top). The results suggest that POF provides more habitat than AMF and MMF for most of the discharges analyzed. There is not much change being observed in shallow-fast habitat. For shallow-slow habitat, there is an increase of habitat area for withdrawal rates below 20 MGD, but there is no change after the rate goes above 20 MGD. Deep-fast habitat exponentially decreases for each flow regime as withdrawal rates increase with POF giving the smallest decay.

## Spectral Analysis in Time Series

We used available USGS discharge daily time series data (gage 02213000) for 1894 – 1909, 1930 – 1974 and 1975 – 2016 that correspond to Pre-dam, Pre-development and Post-development periods, respectively. Using our habitat model, we estimated habitat areas that correspond to the daily mean discharge data. We then performed Spectral Analysis of the generated habitat time series in order to estimate the spectral density that would explain variation in time series. This allowed us to look at the time series in the frequency domain.

**Period before the Lloyd Shoals Dam was built in** 1910**: Pre-dam period** (1894 – 1909)

Figure 5.11 plots daily median values of substrata/structure dependent habitat (as an example) along with a long-term average and a range between daily minimum and daily maximum values during 1894 – 1909 years. It is noticeable, that there is a high variation in the area available for substrata/structure dependent habitat (from 20 to more than 200 acres). However, the variation is changing throughout the year leaving some months (particularly, May and October) with the lowest variation when the available habitat on a particular day was around the median values each year.

Figure 5.12 provides three plots: the top plot is representing time series data of the estimated substrata/structure dependent habitat; the middle and bottom plots represent periodograms obtained using non-parametric and parametric methods, respectively. The time series data is very noisy and does not show any particular trend. The two periodograms suggest that return periods of greater than 100 days have greater power rather than monthly and weekly periods.

**Period after the Lloyd Shoals Dam was built and before** 1975**: Pre-development period** (1930 – 1974)

After the Lloyd Shoals Dam was built in 1910, USGS has missing flow discharge data from August 1912 to September 1929. Therefore, we considered the pre-development period as 1930 – 1974. Figure 5.13 suggests that the variability in substrata/structure dependent habitat dramatically increased during winter and spring seasons; whereas the variability in habitat substantially decreased for late summer and fall seasons (August-November). However, there is not a significance change in mean or median values compared to the pre-dam period 1894 – 1909.

Spectral Analysis for the time series of the Pre-development period suggests that the power of spectral density function increased for return periods greater than 100 days (frequencies lower than 0.01). Figure 5.14 provides smoothed periodograms where the estimated spectral density functions decays at much greater rate when compared to the peridograms from earlier years (1984 – 1909). This means that the sinusoids (from the spectral decomposition) are higher in magnitude for seasonal and annual return periods.

**Period after the Lloyd Shoals Dam was built and after 1975: Post-development period (1975 – 2016)**

During the Post-development time frame, the variability in area suitable substrata/structure dependent habitat is almost constant throughout the year. However, the frequency is noticeably different when we compare a period from January to June (lower frequncy) with a period from July to December (higher frequency). Additionally, during this period the median daily values of available habitat were lower than the long-term average based on the past 41 years (Figure 5.15).

Based on Periodograms given in Figure 5.16, the spectral density function has much higher power at long return period (more than 100 days). Similarly to Figure 5.14 from the second period (1930 – 1974), the sinusoids are higher in magnitude and wider over time when compared to periodograms for the first period (1894 – 1909). For better comparison of the Periodograms from different periods, we plot the smoothed peridograms from parametric approach for the three periods in one graph for each habitat type.

**Comparison between Periodograms from Pre-dam, Pre-development and Post-development periods for each habitat type**

Figure 5.17 clearly shows the shift in spectral density function based on time period for substrata and structure dependent habitat. As we noticed earlier, all periodograms have high powers at seasonal and annual return periods (frequency from 0.003 to 0.01); and have exponential decay for monthly and weekly return periods. The Post-development periodogram has the strongest spectrum at the beginning followed by the Pre-development periodogram. The spectral density function decays almost at the same when we compare periodograms from Pre- and Post- development times. The periodogram from 1894 − 1909 starts at lower magnitude of spectrum but decays at much slower rate.

We also compared the periodograms for other fish habitat types. Figures 5.18− 5.18 represent spectral analysis results for deef-fast, shallow-fast and shallow-slow habitat types, respectively.

The presence of the dam and dam releases gave increased flows on a near roughly annually due to dam releases in the spring. This gave much more deep-slow (structure dependent) flow regime with seasonal and annual return period corresponding to the higher releases in the early spring. The dam smoothed some minor trends at once and twice per month. High spectrum at low frequencies suggests that habitat in Middle Ocmulgee River changes seasonally and annually. This might be important for low cost operations on water management decision based on seasonal and annual patterns.

## 5.4 Conclusion

We present a research framework for water management with focus on ecological impacts and economic development. We examined two reaches and performed hydraulic and habitat modeling. We also applied spectral analysis for time series of four habitat types near Macon. One can imply that the habitat time series would be similar to the Warner Robbins area given similarity in habitat analysis data. Other tools such as the recurrent neural network modeling of the time series could be performed if fish data were available. Issues present in the river that were not modeled include alligators as their presence is temperature driven (Lance, 2003, Joanen & McNease, 1989, Seebacher *et al.*, 2003). In this work, we used an assumption of the channel bathymetry to be unchanging over the years. Ideally, it would be desirable to determine water withdrawals thresholds based on a percent-of-flow basis instead or annual or monthly minimums). Fish species are present that can use each of the flow regimes, thus the model suggests good and not-so-good places to fish for particular species "x" (yet to be tested). The environmental flow concept is much more informative for regulators and the public compared to $7Q10$, because the former carries ecological information. Our future work will include modeling and analysis on the lower reach in Hawkinsville.

## 5.5 Acknowledgements

We would like to thank Michelle Elliott, operation coordinators and Archway professional - Hawkinsville / Pulaski County, for her insightful feedback and support of the Ocmulgee Water Trail Initiative. We would also like to thank Georgia Water Resource Institute (GWRI) for the opportunity to participate in the State Water Resources Research Institute (WRII) Program supported by U.S. Geological Survey (USGS) and for supporting this research. We thank Kenneth W. Swinson, Justin Vale and Kallie Craft for taking part in this research. We also thank Dr. Susan Wilde and research staff at the Warnell School of Forestry and Natural Resources, University of Georgia (UGA) for their assistance in providing sonar equipment for bathymetry data collection.

Table 5.1: Descriptive Statistics (stream flow data 1894 – 1909 : USGS gage 02213000).

| Parameter (units) | Percentile | | | | | |
|---|---|---|---|---|---|---|
| | $0^{th}$ | $25^{th}$ | $50^{th}$ | Mean | $75^{th}$ | $100^{th}$ |
| Stream flow (cfs) | 250 | 1,150 | 1,890 | 2,857 | 3,060 | 53,300 |

Table 5.2: Site Evaluation Criteria used for geospatial anaysis to identify potential landings along the Ocmulgee Water Trail (Macon-Hawkinsville).

| | |
|---|---|
| 1 | paddling distance of 6 – 10 miles |
| 2 | proximity to existing road |
| 3 | contour (elevation and grade) |
| 4 | distance from a population center |
| 5 | river depth |
| 6 | ownership and property licensing |
| 7 | soil type |

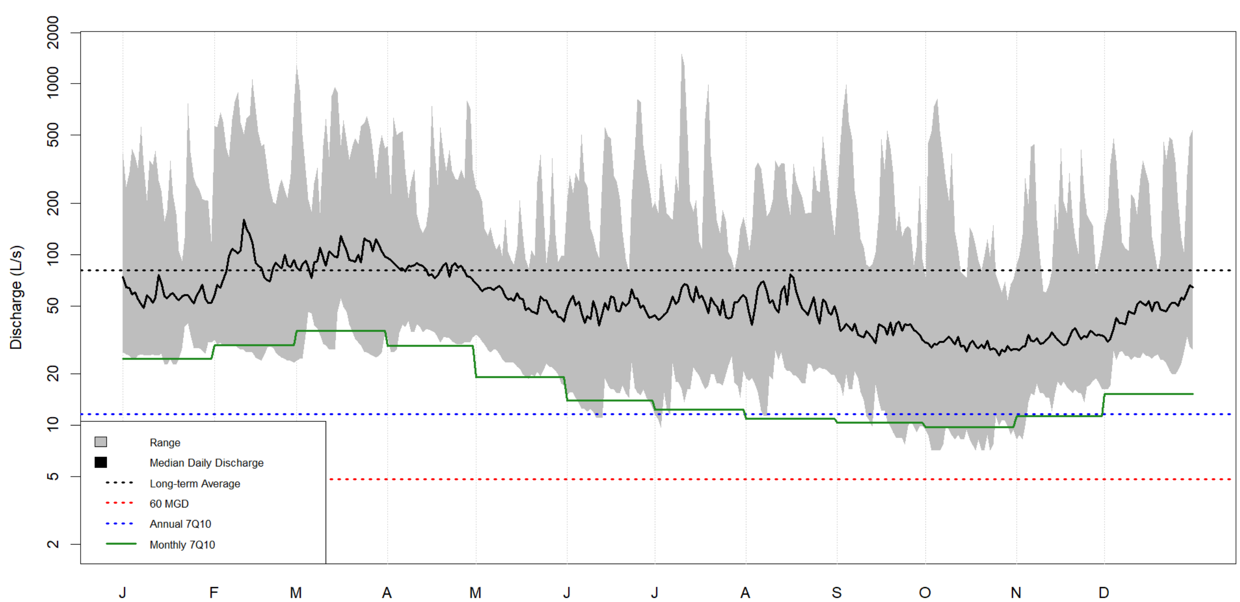



Figure 5.1: Long-term, minimally altered hydrograph on the Ocmulgee River near Macon (1894 – 1909). The shaded area represents the lowest and highest discharge observed on each day of the year, the solid black line is the daily median, the dashed black line is the long-term mean, the red dashed line is the maximum allowed withdrawal, and the blue dashed line is the Annual $7Q10$ level.

Figure 5.2: Bathymetry data collection

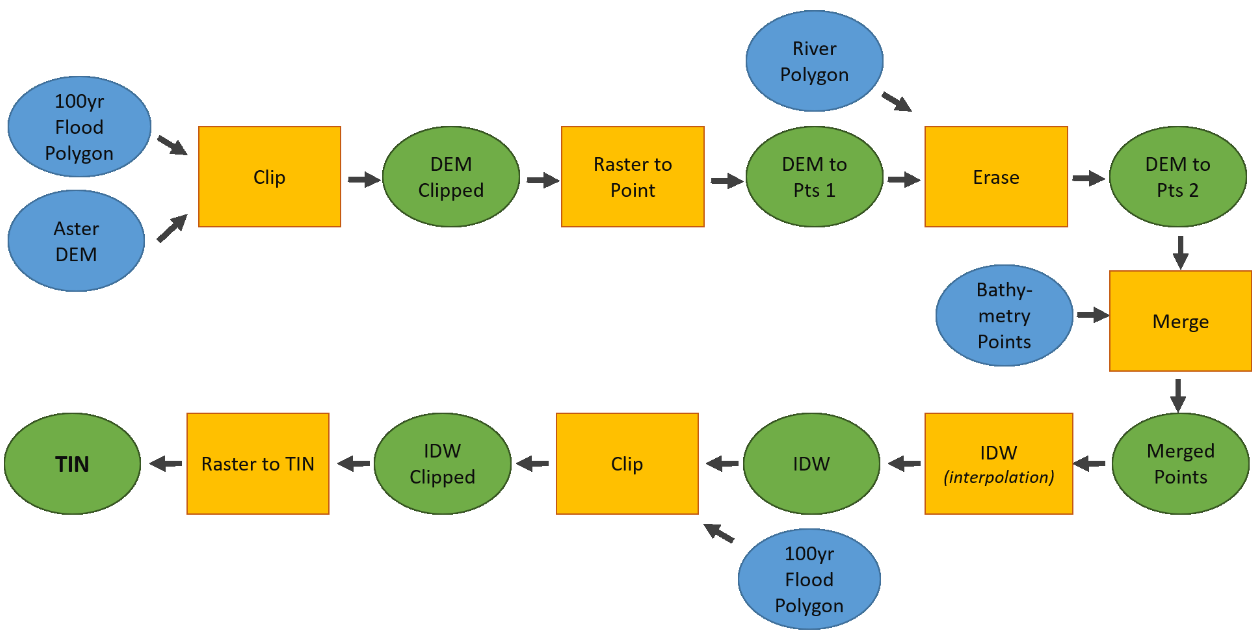



Figure 5.3: ArcMap Model for obtaining TIN Surface.

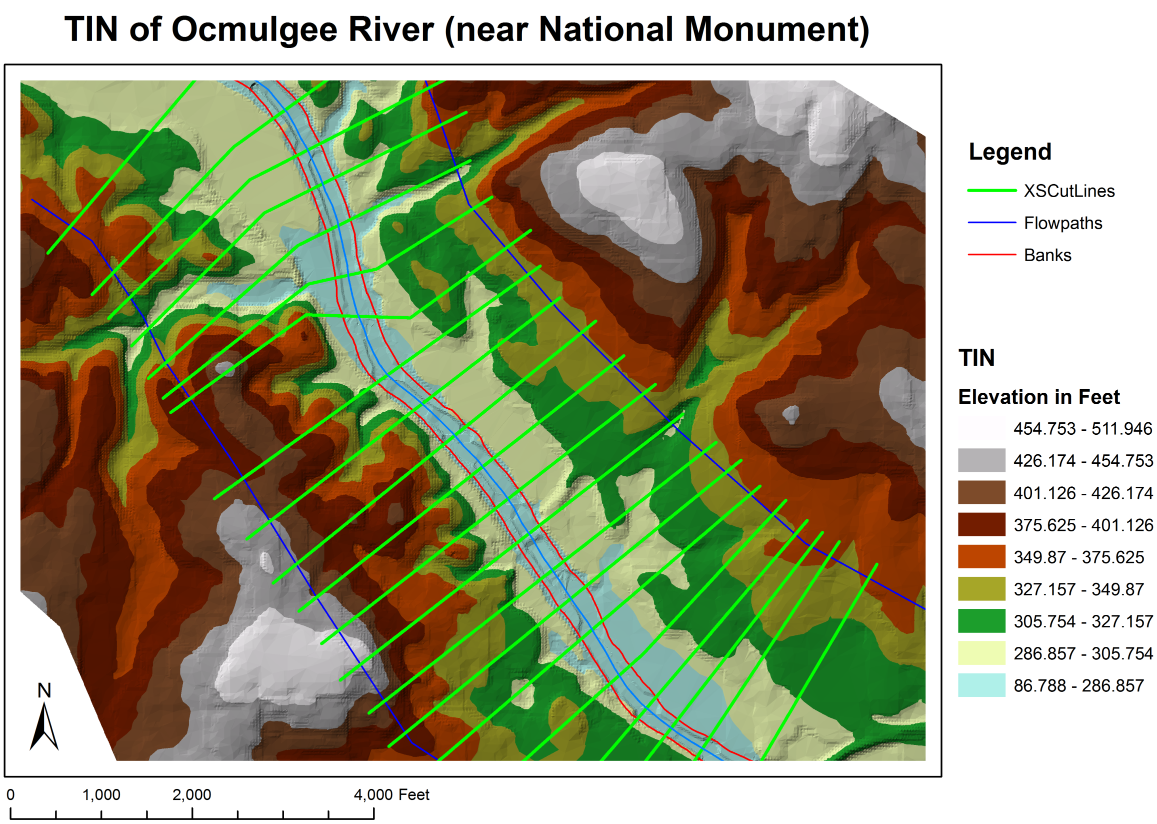



Figure 5.4: Map of the reach on the Ocmulgee River near Macon, Georgia. Surveyed cross-sections are shown in green, general flow paths of the main channel and floodplains as blue lines, and bank demarcation points as red lines.

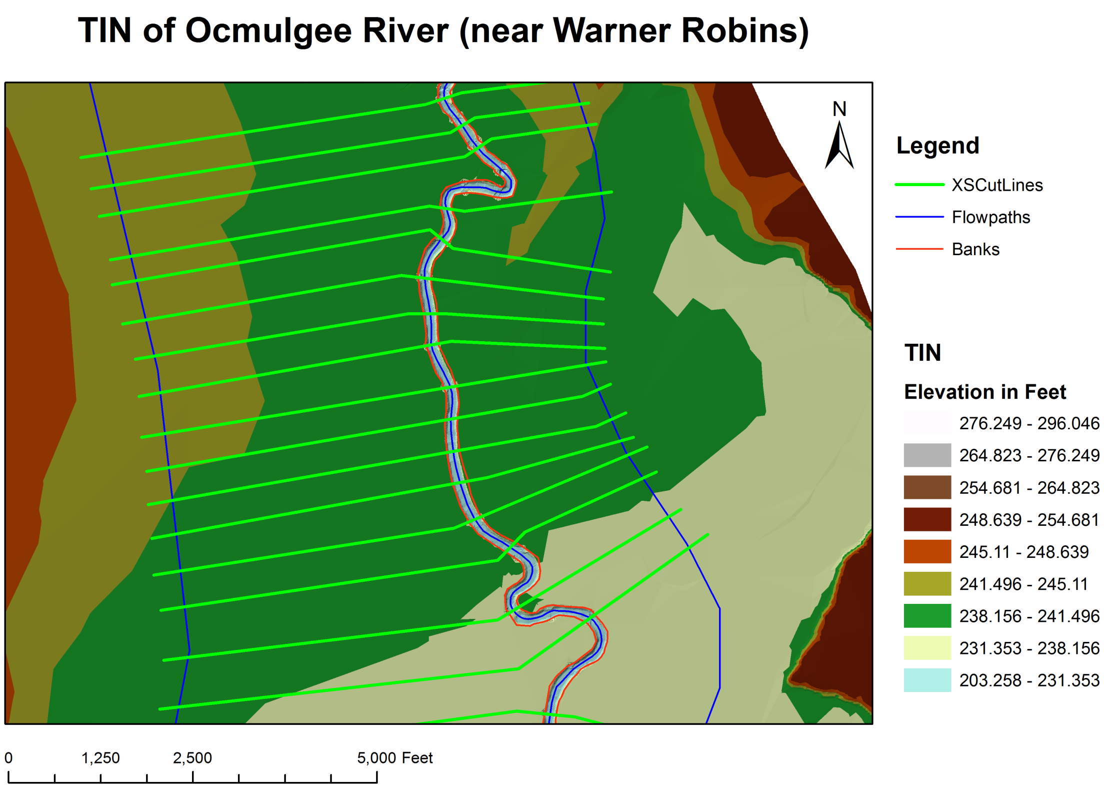



Figure 5.5: Map of the reach on the Ocmulgee River near Warner Robins, Georgia. Surveyed cross-sections are shown in green, general flow paths of the main channel and floodplains as blue lines, and bank demarcation points as red lines.

## Landings along the Ocmulgee River

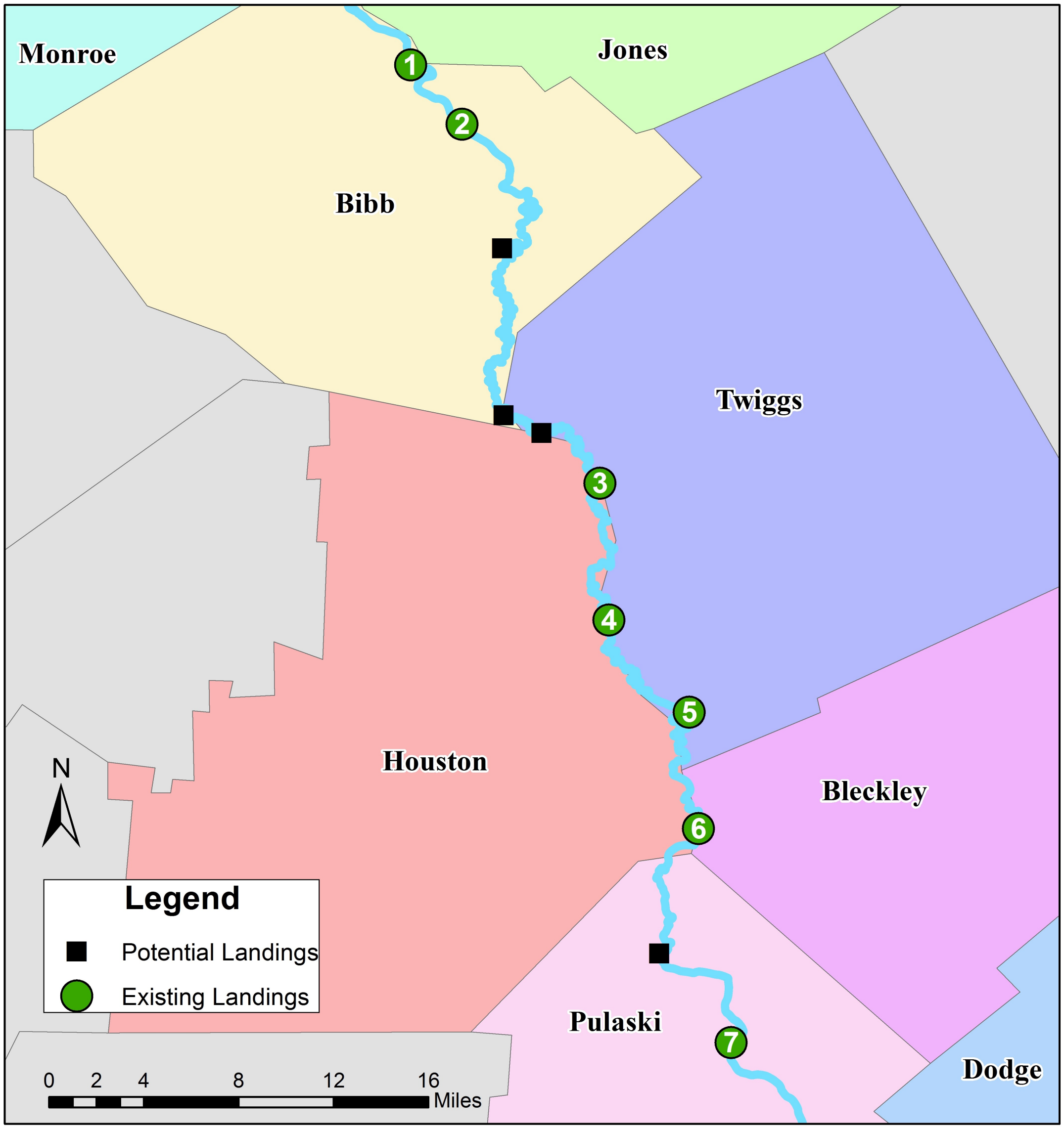



Figure 5.6: The Ocmulgee Water Trail and its existing and potential landings

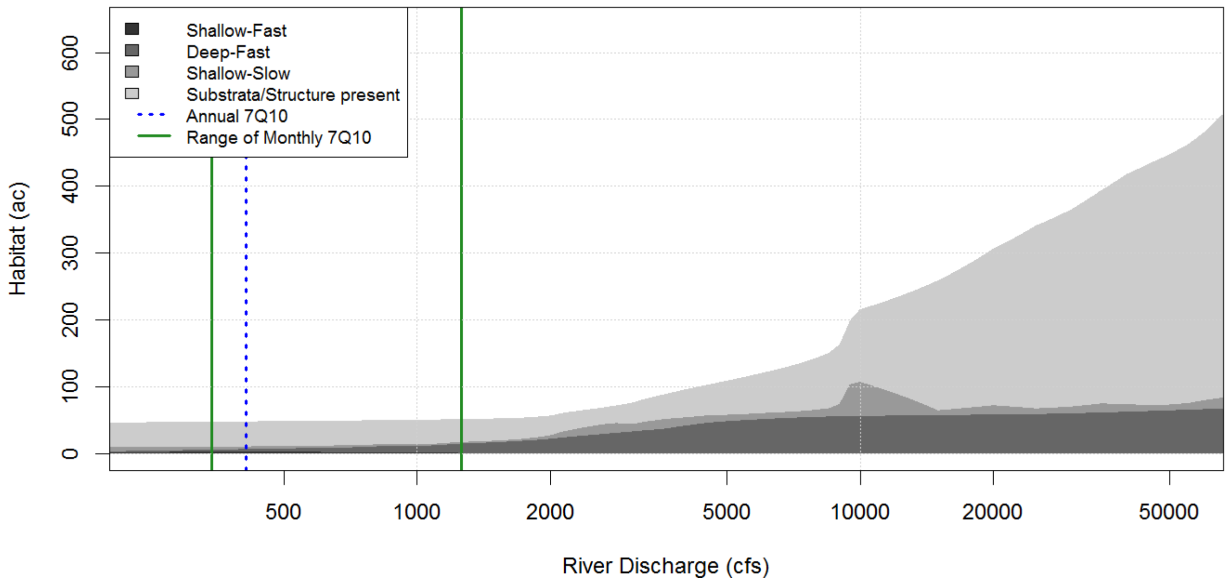



Figure 5.7: Cumulative habitat rating curves over the range of discharges observed in the Ocmulgee river near Macon

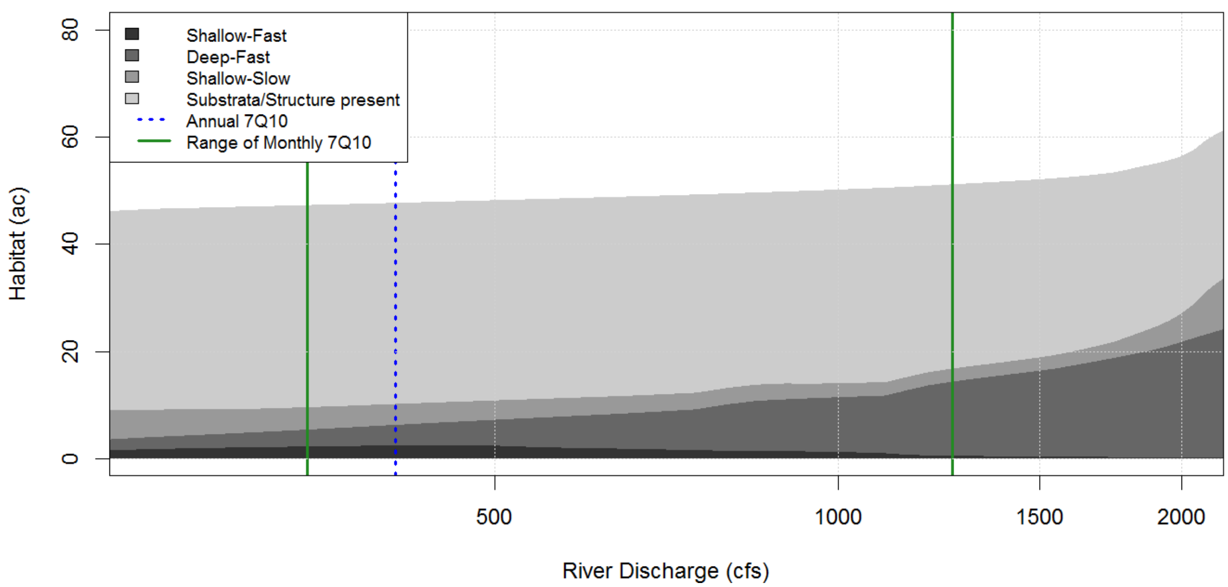



Figure 5.8: Cumulative habitat rating curves over the range of low discharges (250 – 2000 cfs) observed in the Ocmulgee river near Macon

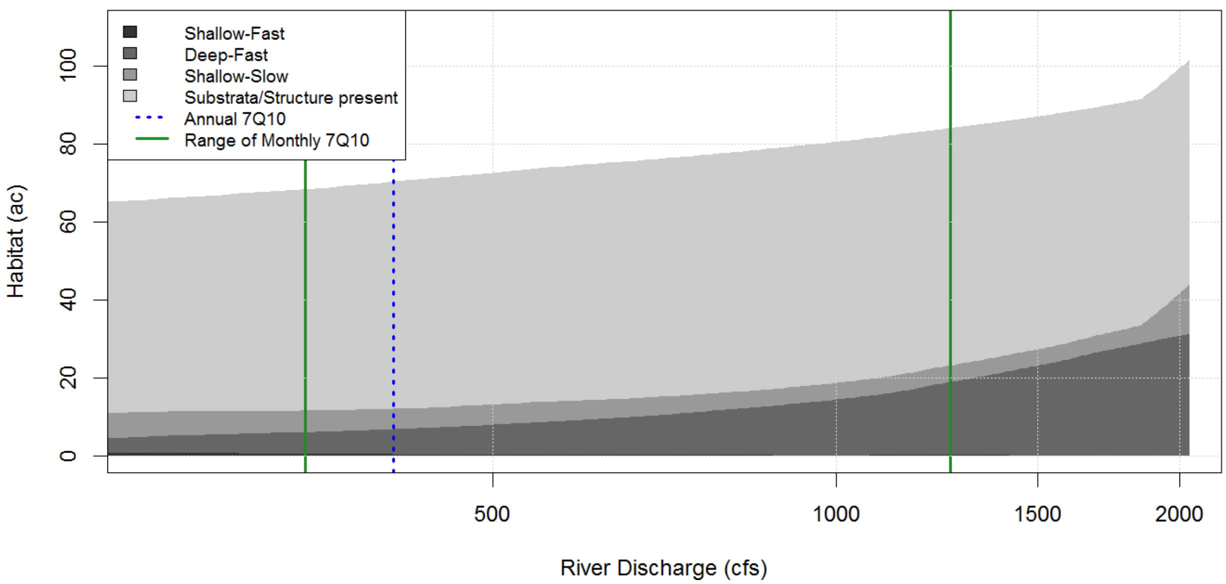



Figure 5.9: Cumulative habitat rating curves over the range of low discharges (250 – 2000 cfs) observed in the Ocmulgee river near Warner Robins

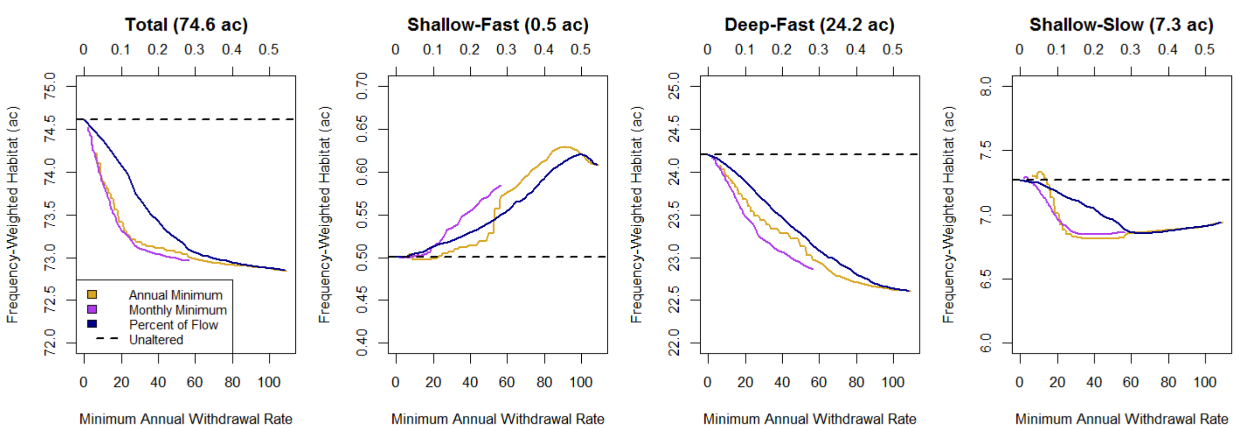



Figure 5.10: Habitat computed as a frequency-weighted quantity using effectiveness analysis: comparison of environmental flow alternatives

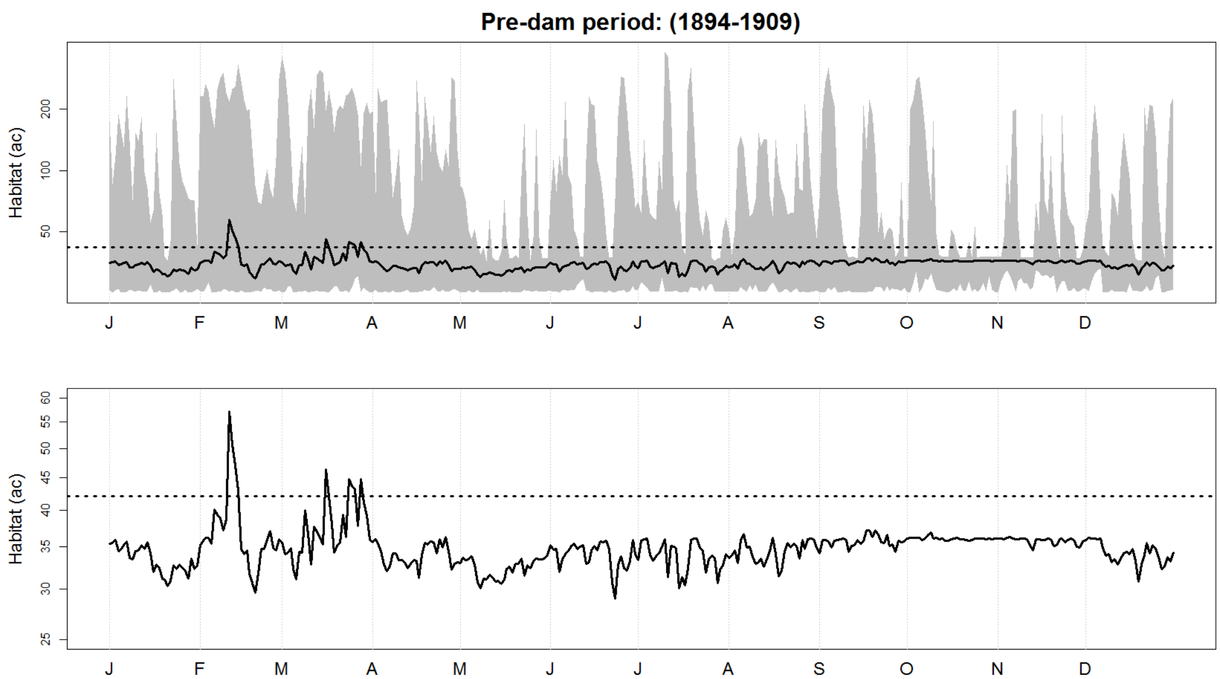



Figure 5.11: Time Series of substrata/structure dependent habitat near Macon during Pre-dam period (1894 − 1909). The shaded area represents the lowest and highest habitat area on each day of the year, the solid black line is the daily median, the dashed black line is the long-term mean.

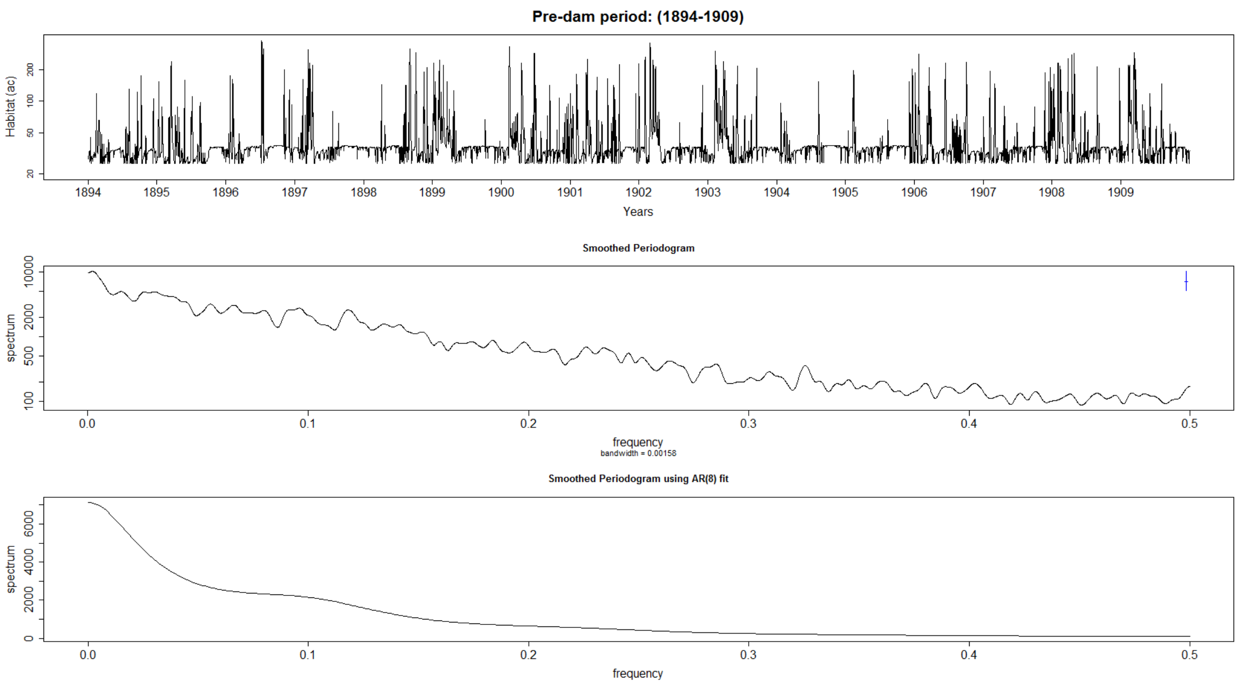



Figure 5.12: Spectral Analysis: (top) Time Series of substrata/structure dependent habitat near Macon during 1894 – 1909; (middle) Smoothed Periodogram using non-parametric approach; (bottom) Smoothed Periodogram using parametric approach

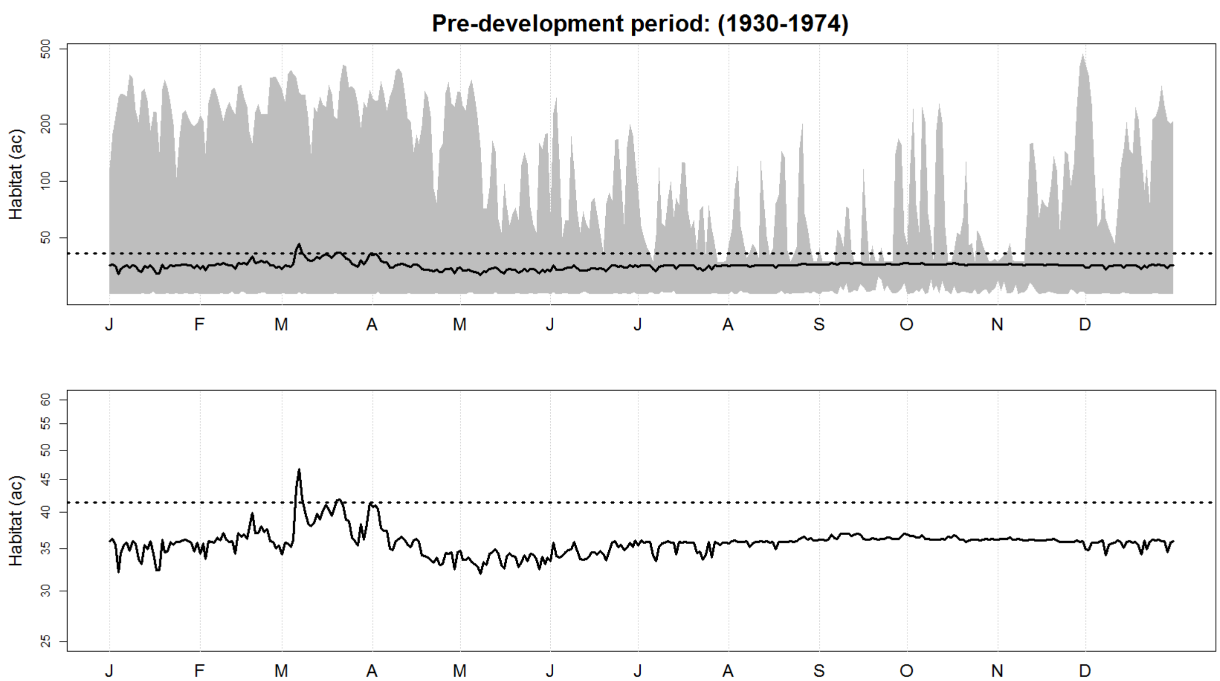



Figure 5.13: Time Series of substrata/structure dependent habitat near Macon during Pre-development period (1930 – 1974). The shaded area represents the lowest and highest habitat area on each day of the year, the solid black line is the daily median, the dashed black line is the long-term mean.

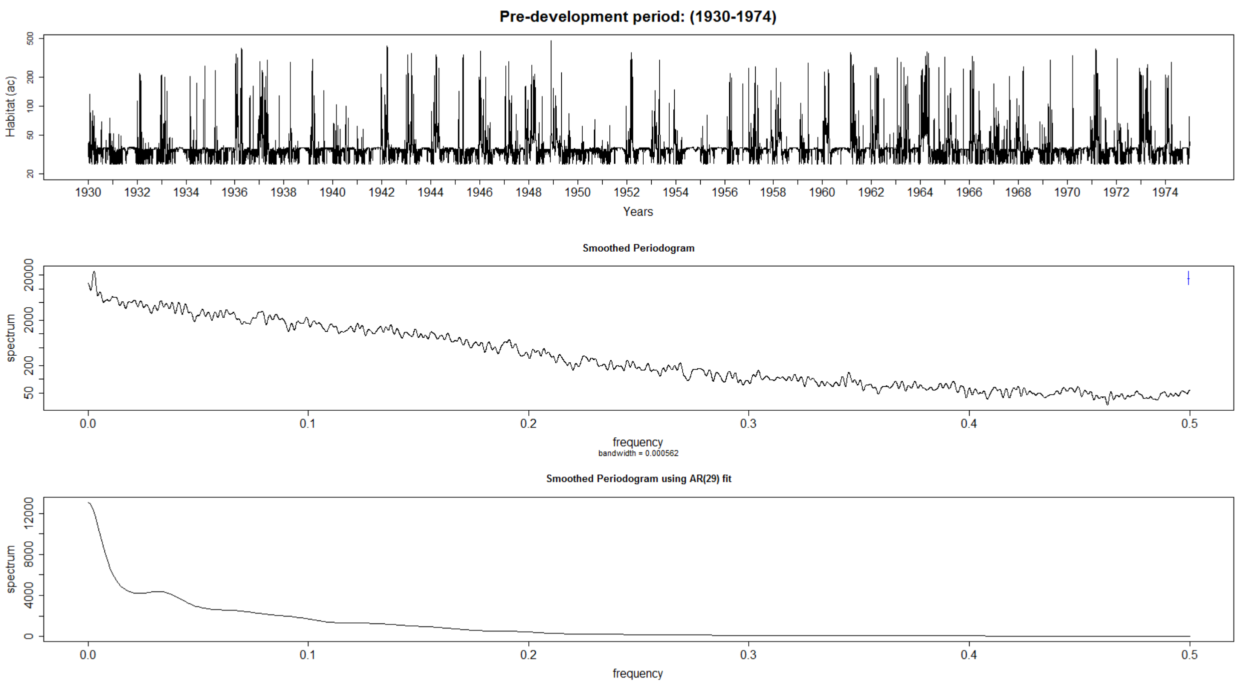



Figure 5.14: Spectral Analysis: (top) Time Series of substrata/structure dependent habitat near Macon during Pre-development period (1930 – 1974); (middle) Smoothed Periodogram using non-parametric approach; (bottom) Smoothed Periodogram using parametric approach

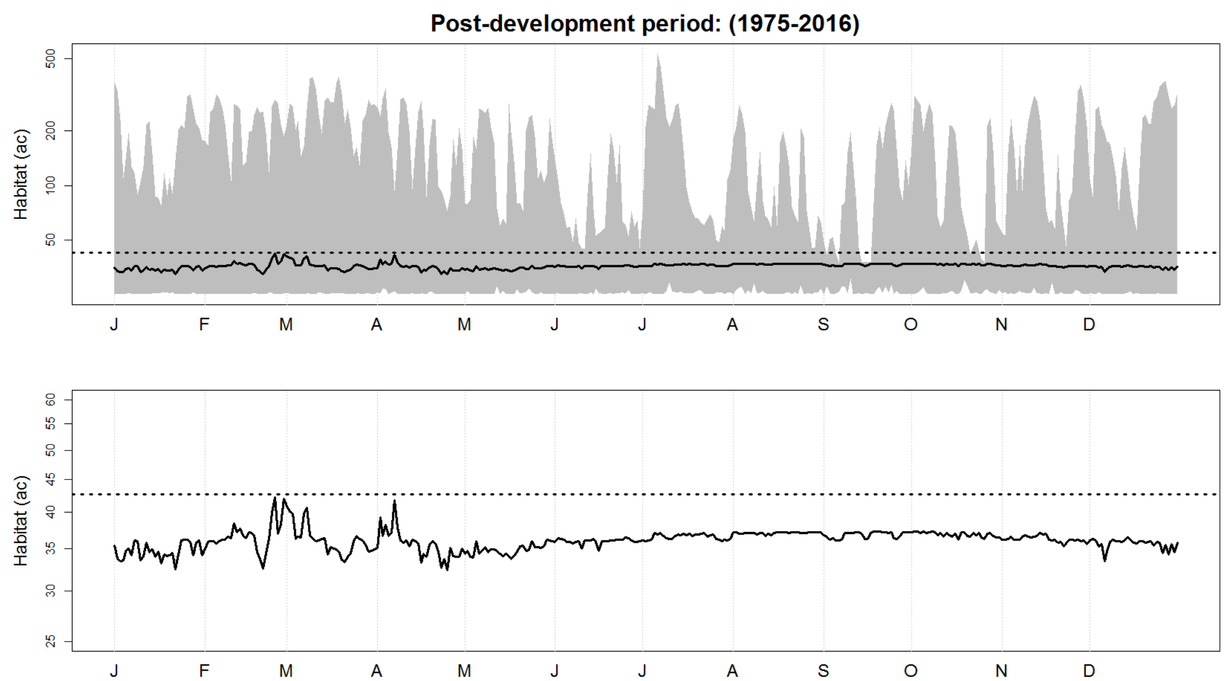



Figure 5.15: Time Series of substrata/structure dependent habitat near Macon during Post-development period (1975 – 2016). The shaded area represents the lowest and highest habitat area on each day of the year, the solid black line is the daily median, the dashed black line is the long-term mean.

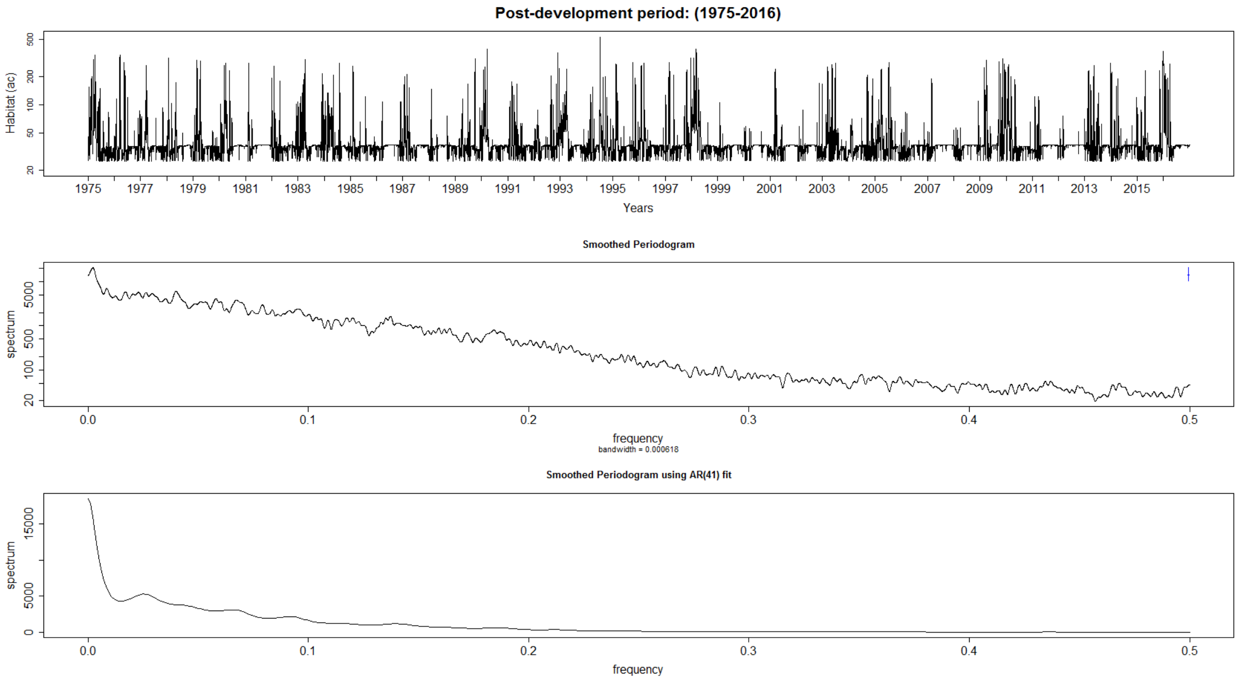



Figure 5.16: Spectral Analysis: (top) Time Series of substrata/structure dependent habitat near Macon during Post-development period (1975 – 2016); (middle) Smoothed Periodogram using non-parametric approach; (bottom) Smoothed Periodogram using parametric approach

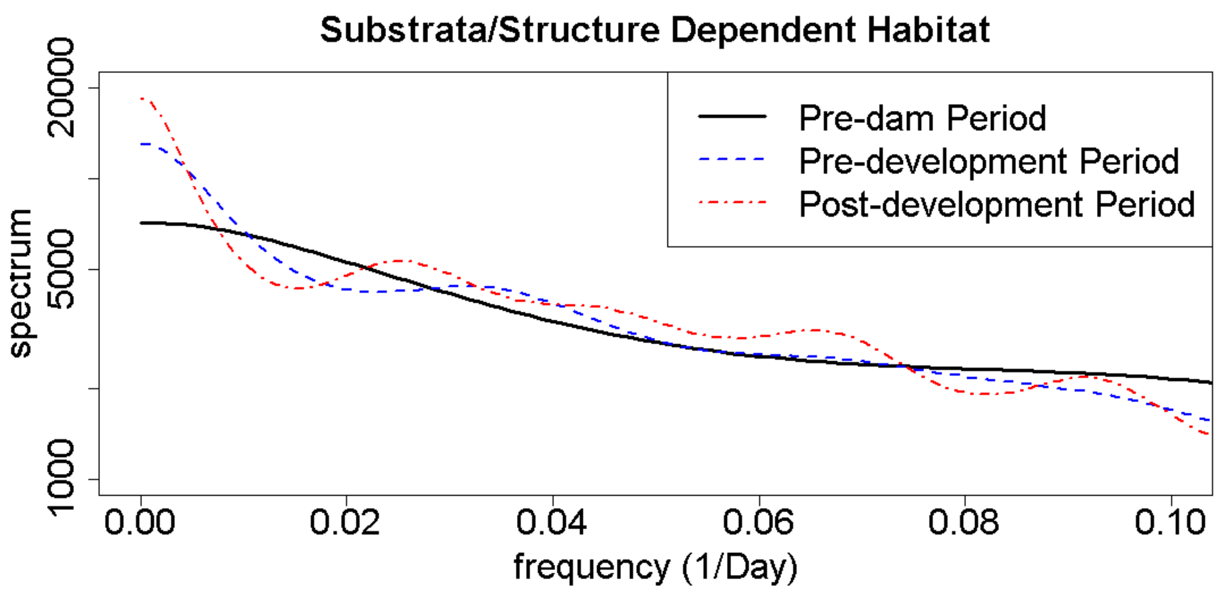



Figure 5.17: Spectral Analysis: Comparison of Smoothed Periodogram using AutoRegressive function between three periods for substrata/structure dependent habitat

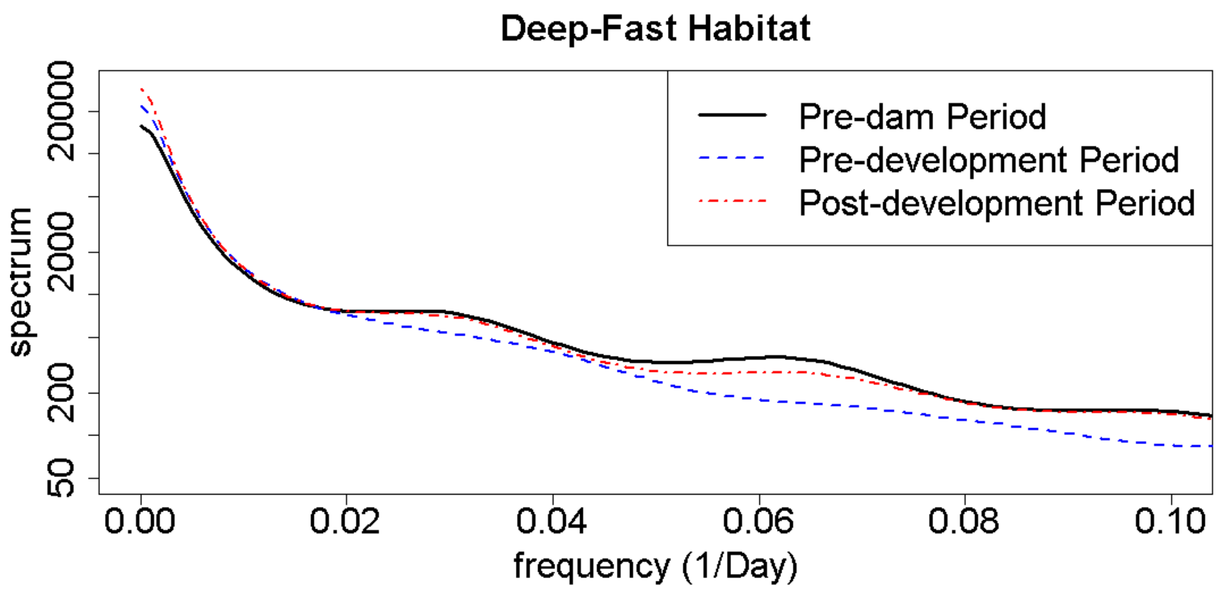



Figure 5.18: Spectral Analysis: Comparison of Smoothed Periodogram using AutoRegressive function between three periods for Deep-Fast habitat

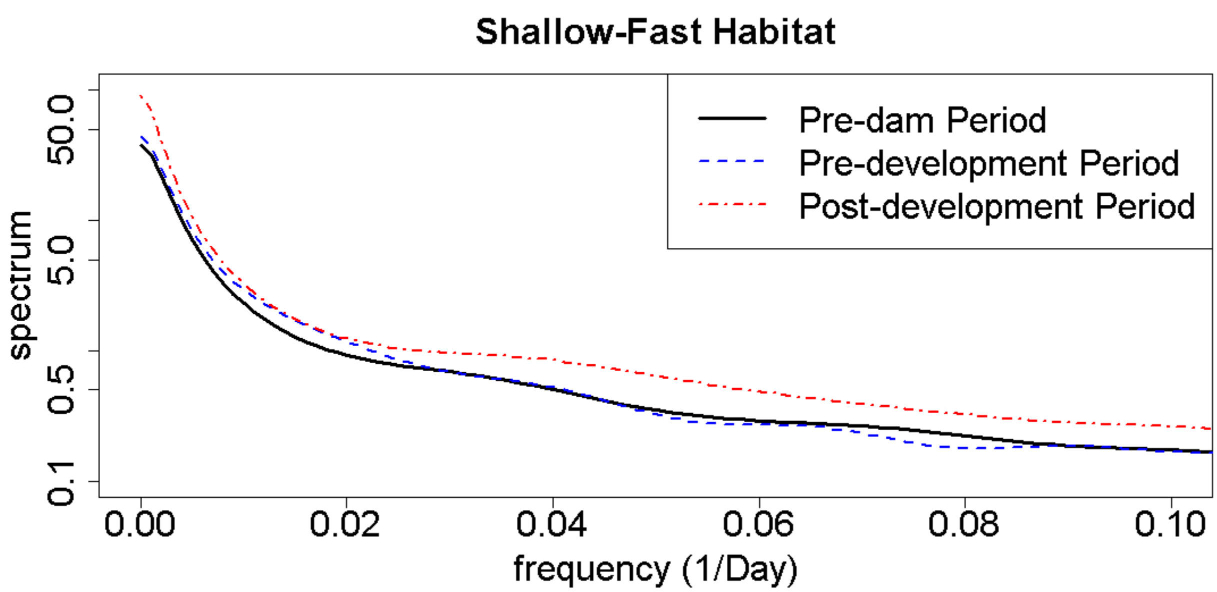



Figure 5.19: Spectral Analysis: Comparison of Smoothed Periodogram using AutoRegressive function between three periods for Shallow-Fast habitat

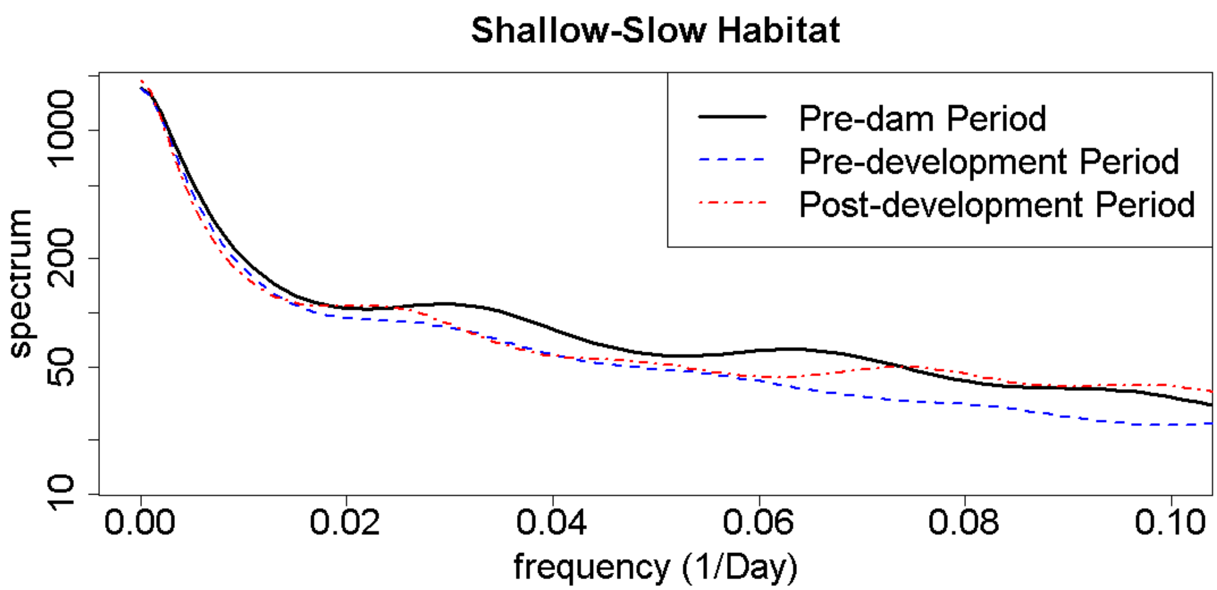



Figure 5.20: Spectral Analysis: Comparison of Smoothed Periodogram using AutoRegressive function between three periods for Shallow-Slow habitat

# Chapter 6

# Recommendation, Conclusion and Future Research

## 6.1 Recommendations

For windrow composting pad management, keeping waste volume at low levels and pond volume at medium levels is preferred if the operators of windrow composting system to maintain TSS and $NO_3$ levels. Additionally, operators can prevent high BOD levels by keeping waste volume at low levels during winter and summer months. Spraying the water from the pond on the composting pad instead of discharging the effluent during summer and early fall months. Also, having a pond with sufficient capacity to maintain a four to six foot depth of operational storage while having additional storage for an extreme event storm (a 25 or 50 year storm) to suppresses biochemical activity in the pond.

For water management of the Middle Oconee and the Ocmulgee Rivers, it would be desirable to determine water withdrawals thresholds based on a percent-of-flow basis instead of annual or monthly minimums. The environmental flow concept is much more informative for regulators and the public compared to $7Q10$, because the former carries ecological information. Fish habitat types are present that can use each of the flow regimes, thus the model suggests good and not-so-good places to fish for species “x” (yet to be tested). Additionally, the operation cost can be reduced as periodograms for all four fish habitat types suggest seasonal and annual habitat changes based on spectral analysis results.

## 6.2 Conclusion

This dissertation developed integrated solutions for hydrological, water quality and ecological applications using hydraulic modeling as well as spatial and time series analysis. The work focused on building a foundation and developing new knowledge for the improvement of water resource management.

We demonstrated significance of recurrent neural network’s sensitivity analysis to overcome the black-box nature of the artificial neural network and gain insight into the hydrological system. The modeling approach of recurrent neural network can be further applied in other studies of dynamic ecological systems because it generalizes in a straightforward manner to nearly any scenario. Olden’s algorithm and Lek’s profile method showed that Pond Volume, Waste Volume and Temperature have maximum effect on Total Suspended Solids (TSS) concentration whereas Temperature and Waste Volume are the most important variables for Biological Oxygen Demand (BOD) prediction.

Two case studies of water resource management in the Ocmulgee and the Middle Oconee Rivers, demonstrated a basis for examining trade-offs in water management between ecological impacts and economic development in order to meet future demands in Georgia. Additionally, we considered windrow composting pad management in order to enable operators to anticipate conditions when water quality concentration exceeds regulatory limits. Time series analysis and physics-based modeling approaches allowed us to improve model performance and gain more insights into the dynamics of the system. The detailed sensitivity analysis of a data-driven methods such as a recurrent neural network allowed a better understanding of water quality dynamics of collected runoff and assist in identifying strategies for better management of windrow composting systems. Similar analysis would be possible with various fish species if fish population data were available. The Spectral Analysis for time series (frequency domain) allowed us to explain variation in time series based on historic events using both non-parametric and parametric approaches. A comprehensive analysis of environmental flow schemes represented a key step in ensuring adequate water availability to meet increasing human needs while minimizing adverse impacts on aquatic ecosystems.

## 6.3 Future Research

Future work would incorporate water management and habitat analysis on the lower part of the Ocmulgee River near Hawkinsville. We will also look at hourly temporal resolution and compare the results with the daily temporal resolution (used in this study). If fish data become available, we will build recurrent neural network for expanding our analysis for other locations along the Middle Ocmulgee River. Additionally, we will include integration of hydrological, hydraulic and ecological models such as HEC-RAS, SWAT and habitat suitability models to evaluate ecological responses to hydrological and land use changes. The Soil and Water Assessment Tool (SWAT) hydrologic model would allow us to predict the rate of runoff and sediment yield in a watershed by taking into account soil, weather, digital elevation and land use data.

# References


Araghinejad, Shahab. 2014. *Data-Driven Modeling: Using MATLAB in Water Resources and Environmental Engineering*. Vol. 67. Springer.

Arthington, Angela H, Bunn, Stuart E, Poff, N LeRoy, & Naiman, Robert J. 2006. The challenge of providing environmental flow rules to sustain river ecosystems. *Ecological Applications*, **16**(4), 1311–1318.

Bain, MB. 1995. Habitat at the local scale: multivariate patterns for stream fishes. *Bulletin Francais de la Peche et de la Pisciculture (France)*.

Beale, Mark Hudson, Hagan, Martin T, & Demuth, Howard B. 2012. Neural network toolboxâĎć userâĂŹs guide. *In: R2012a, The MathWorks, Inc., 3 Apple Hill Drive Natick, MA 01760-2098, www. mathworks. com*. Citeseer.

Beck, Marcus. 2015. NeuralNetTools: Visualization and Analysis Tools for Neural Networks. *R package version*, **1**(0).

Bhattacharjee, Natalia V, R, Willis Joshua, W, Tollner Ernest, McKay, & Kyle, S. 2017. *Habitat provision associated with environmental flows*. Tech. rept. In review by U.S. Army Engineer Research and Development Center.

Bowen, Zachary H, Freeman, Mary C, & Bovee, Ken D. 1998. Evaluation of generalized habitat criteria for assessing impacts of altered flow regimes on warmwater fishes. *Transactions of the American Fisheries Society*, **127**(3), 455–468.

Brunner, Gary W. 2001. *HEC-RAS River Analysis System: User's Manual*. US Army Corps of Engineers, Institute for Water Resources, Hydrologic Engineering Center.

Campana, Pete, Knox, John, Grundstein, Andrew, & Dowd, John. 2012. The 2007-2009 Drought in Athens, Georgia, United States: A Climatological Analysis and an Assessment of Future Water Availability1. *JAWRA Journal of the American Water Resources Association*, **48**(2), 379–390.

Chen, Assaf, Abramson, Adam, Becker, Nir, & Megdal, Sharon B. 2015. A tale of two rivers: Pathways for improving water management in the Jordan and Colorado River basins. *Journal of Arid Environments*, **112**, 109–123.

Chow, Te Ven. 1959. *Open channel hydraulics*. McGraw-Hill Book Company, Inc; New York.

Commision, Middle Georgia Regional. 2012. *Multi-Region River Corridor Feasibility Study: Phase II. Middle Georgia Region*.

Comum, WCED O Nosso Futuro. 1987. World Commision on Environment and Development. *Meribérica/Liber Editores: Lisboa, Portugal*.

Dale, Virginia H. 2003a. Opportunities for using ecological models for resource management. *Pages 3–19 of: Ecological Modeling for Resource Management*. Springer.

Dale, Virginia H. 2003b. New directions in ecological modeling for resource management. *Pages 310–320 of: Ecological Modeling for Resource Management*. Springer.

Declaration, Brisbane. 2007. The Brisbane Declaration: environmental flows are essential for freshwater ecosystem health and human well-being. *Pages 3–6 of: 10th International River Symposium, Brisbane, Australia*.

DeCoursey, Donn G. 1966. runoff hydrograph equation.

Dillon, Jeff. 2013. Comparison of two rivers i northern Michigan for determination of differential effects of development.

Dorahy, CG, Pirie, AD, McMaster, I, Muirhead, L, Pengelly, P, Chan, KY, Jackson, M, & Barchia, IM. 2009. Environmental Risk Assessment of Compost Prepared from Salvinia, Egeria densa, and Alligator Weed. *Journal of environmental quality*, **38**(4), 1483–1492.

Doyle, Martin W, Stanley, Emily H, Strayer, David L, Jacobson, Robert B, & Schmidt, John C. 2005. Effective discharge analysis of ecological processes in streams. *Water Resources Research*, **41**(11).

Dudgeon, David, Arthington, Angela H, Gessner, Mark O, Kawabata, Zen-Ichiro, Knowler, Duncan J, Lévêque, Christian, Naiman, Robert J, Prieur-Richard, Anne-Hélène, Soto, Doris, Stiassny, Melanie LJ, *et al.* 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological reviews*, **81**(2), 163–182.

Duncan, OJ, Tollner, EW, Ssegane, H, & McCutcheon, SC. 2013a. Curve Number Approaches to Estimate Drainage from a Yard Waste Windrow Composting Pad. *Applied Engineering in Agriculture*, **29**(2), 201–208.

Duncan, OJ, Tollner, EW, & Ssegane, H. 2013b. An Instantaneous Unit Hydrograph for Estimating Runoff from Windrow Composting Pads. *Applied Engineering in Agriculture*, **29**(2), 209–223.

Eberhart, Russell C, & Shi, Yuhui. 2007. Computational Intelligence: concepts to implementation. *Imprint: Morgan Kaufmann, Elsevier, ISBN*, **1918074598**, 269–372.

Fisher, Donna, & Thompson, Ben. 2003. Basin Water Plans for GeorgiaâĂŹs Coastal Region: The. *Empty ShelfâĂİ of Data Critical for the Planning Process Water Policy Working Paper*, **3**.

Freeman, Mary C, Bowen, Zachary H, & Crance, Johnie H. 1997. Transferability of habitat suitability criteria for fishes in warmwater streams. *North American Journal of Fisheries Management*, **17**(1), 20–31.

Garbrecht, Jurgen D. 2006. Comparison of three alternative ANN designs for monthly rainfall-runoff simulation. *Journal of Hydrologic Engineering*, **11**(5), 502–505.

Garson, David G. 1991. Interpreting neural network connection weights.

Georgakakos, AP, Yao, H, Kistenmacher, M, Georgakakos, KP, Graham, NE, Cheng, F-Y, Spencer, C, & Shamir, E. 2012. Value of adaptive water resources management in Northern California under climatic variability and change: Reservoir management. *Journal of hydrology*, **412**, 34–46.

Gevrey, Muriel, Dimopoulos, Ioannis, & Lek, Sovan. 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological modelling*, **160**(3), 249–264.

Giam, Xingli, & Olden, Julian D. 2015. A new R 2-based metric to shed greater insight on variable importance in artificial neural networks. *Ecological Modelling*, **313**, 307–313.

Gibson, CA, Meyer, JL, Poff, NL, Hay, LE, & Georgakakos, A. 2005. Flow regime alterations under changing climate in two river basins: implications for freshwater ecosystems. *River Research and Applications*, **21**(8), 849–864.

Goh, ATC. 1995. Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, **9**(3), 143–151.

Gotvald, Anthony J. 2016. *Selected low-flow frequency statistics for continuous-record streamgages in Georgia, 2013*. Tech. rept. US Geological Survey.

Gregory, S, Wildman, R, Ashkenas, L, Wildman, K, & Haggerty, P. 2002. Fish assemblages. Willamette River Basin Planning Atlas: Trajectories of Environmental and Ecological Change. *Oregon State University Press, Corvallis, Oregon*.

Grubaugh, Jack W, & Wallace, J Bwce. 1995. Functional structure and production of the benthic community in a Piedmont river: 1956-1957 and 1991-1992. *Limnology and Oceanography*, **40**(3), 490–501.

Haan, Charles Thomas. 2002. *Statistical methods in hydrology*. The Iowa State University Press.

Hickey, J, & Fields, W. 2009. HEC-EFM Ecosystem Functions Model: Quick Start Guide. *US Army Corps of Engineers, Institute for Water Resources, Hydrologic Engineering Center*.

Huggins, LF, & Burney, JR. 1982. Surface runoff, storage and routing. *Hydrologic modeling of small watersheds*, 169–225.

Jackson, Robert B, Carpenter, Stephen R, Dahm, Clifford N, McKnight, Diane M, Naiman, Robert J, Postel, Sandra L, & Running, Steven W. 2001. Water in a changing world. *Ecological applications*, **11**(4), 1027–1045.

Jeong, Kwang-Seuk, Joo, Gea-Jae, Kim, Hyun-Woo, Ha, Kyong, & Recknagel, Friedrich. 2001. Prediction and elucidation of phytoplankton dynamics in the Nakdong River (Korea) by means of a recurrent artificial neural network. *Ecological Modelling*, **146**(1), 115–129.

Jeong, Kwang-Seuk, Kim, Dong-Kyun, & Joo, Gea-Jae. 2006. River phytoplankton prediction model by Artificial Neural Network: Model performance and selection of input variables to predict time-series phytoplankton proliferations in a regulated river system. *Ecological Informatics*, **1**(3), 235–245.

Joanen, Ted, & McNease, Larry L. 1989. Ecology and physiology of nesting and early development of the American alligator. *American Zoologist*, **29**(3), 987–998.

Jones, NE. 2014. The dual nature of hydropeaking rivers: Is ecopeaking possible? *River Research and Applications*, **30**(4), 521–526.

Jones-Farrand, D Todd, Fearer, Todd M, Thogmartin, Wayne E, III, Frank R Thompson, Nelson, Mark D, & Tirpak, John M. 2011. Comparison of statistical and theoretical habitat models for conservation planning: the benefit of ensemble prediction. *Ecological Applications*, **21**(6), 2269–2282.

Jowett, IG. 1997. Instream flow methods: a comparison of approaches. *Regulated Rivers: Research & Management*, **13**(2), 115–127.

Kalaba, L, Wilson, B G, & Haralampides, K. 2007. A Storm Water Runoff Model For Open Windrow Composting Sites. *Compost Science and Utilization*, **15**(3), 142–150.

Kalaba, Ljubica, & Wilson, Bruce G. 2005. Modeling stormwater runoff from open windrow composting sites. *In: 33rd CSCE Annual Conference, Toronto, ON, Canada, 2-4 June 2005*. Canadian Society for Civil Engineering.

Katz, Rachel A, & Freeman, Mary C. 2015. Evidence of population resistance to extreme low flows in a fluvial-dependent fish species. *Canadian Journal of Fisheries and Aquatic Sciences*, **72**(11), 1776–1787.

Kim, Dong-Kyun, Jeong, Kwang-Seuk, Whigham, Peter A, & Joo, Gea-Jae. 2007. Winter diatom blooms in a regulated river in South Korea: explanations based on evolutionary computation. *Freshwater Biology*, **52**(10), 2021–2041.

Klipsch, Joan D, & Hurst, Marilyn B. 2013. *HEC-ResSim Reservoir System Simulation: User's Manual*. US Army Corps of Engineers, Institute for Water Resources, Hydrologic Engineering Center, Davis, California.

LaFontaine, Jacob H, Hay, Lauren E, Viger, Roland J, Regan, R Steve, & Markstrom, Steven L. 2015. Effects of climate and land cover on hydrology in the southeastern US: potential impacts on watershed planning. *JAWRA Journal of the American Water Resources Association*, **51**(5), 1235–1261.

Lance, Valentine A. 2003. Alligator physiology and life history: the importance of temperature. *Experimental gerontology*, **38**(7), 801–805.

Lane, Stuart N, & Richards, Keith S. 1997. Linking river channel form and process: time, space and causality revisited. *Earth Surface Processes and Landforms*, **22**(3), 249–260.

Lek, Sovan, Delacoste, Marc, Baran, Philippe, Dimopoulos, Ioannis, Lauga, Jacques, & Aulagnier, Stéphane. 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological modelling*, **90**(1), 39–52.

Liu, Qian-Jin, Shi, Zhi-Hua, Fang, Nu-Fang, Zhu, Hua-De, & Ai, Lei. 2013. Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet–ANN approach. *Geomorphology*, **186**, 181–190.

Maier, Holger R, & Dandy, Graeme C. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental modelling & software*, **15**(1), 101–124.

McKay, S Kyle. 2013. *Alternative environmental flow management schemes*. Tech. rept. DTIC Document.

McKay, S Kyle. 2015. Quantifying Tradeoffs Associated with Hydrologic Environmental Flow Methods. *JAWRA Journal of the American Water Resources Association*, **51**(6), 1508–1518.

McKay, S Kyle, Freeman, Mary C, & Covich, Alan P. 2016. Application of Effective Discharge

Analysis to Environmental Flow Decision-Making. *Environmental management*, **57**(6), 1153–1165.

McKay, Steven Kyle. 2014. *Informing Flow Management Decisions in the Middle Oconee River*. Ph.D. thesis, University of Georgia.

Milot, Julie, Rodriguez, Manuel J, & Sérodes, Jean B. 2002. Contribution of neural networks for modeling trihalomethanes occurrence in drinking water. *Journal of water resources planning and management*, **128**(5), 370–376.

Moog, Otto. 1993. Quantification of daily peak hydropower effects on aquatic fauna and management to minimize environmental impacts. *Regulated Rivers: Research & Management*, **8**(1-2), 5–14.

Nelson, Daniel J, & Scott, Donald C. 1962. Role of detritus in the productivity of a rock-outcrop community in a Piedmont stream. *Limnology and Oceanography*, **7**(3), 396–413.

NGRC. 2011. *Northeast Georgia Plan 2035: Regional Assessment*.

Nourani, Vahid, & Fard, Mina Sayyah. 2012. Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. *Advances in Engineering Software*, **47**(1), 127–146.

Olden, Julian D, & Jackson, Donald A. 2002. Illuminating the âĂIJblack boxâĂİ: a randomization approach for understanding variable contributions in artificial neural networks. *Ecological modelling*, **154**(1), 135–150.

Olden, Julian D, Joy, Michael K, & Death, Russell G. 2004. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*, **178**(3), 389–397.

Pisoni, Enrico, Farina, Marcello, Carnevale, Claudio, & Piroddi, Luigi. 2009. Forecasting peak air pollution levels using NARX models. *Engineering Applications of Artificial Intelligence*, **22**(4), 593–602.

Poff, N LeRoy. 2009. *Managing for variability to sustain freshwater ecosystems*.

Poff, N LeRoy, & Allan, J David. 1995. Functional organization of stream fish assemblages in relation to hydrological variability. *Ecology*, **76**(2), 606–627.

Poff, N LeRoy, Allan, J David, Bain, Mark B, Karr, James R, Prestegaard, Karen L, Richter, Brian D, Sparks, Richard E, & Stromberg, Julie C. 1997. The natural flow regime. *BioScience*, 769–784.

Richter, Brian D. 2010. Re-thinking environmental flows: from allocations and reserves to sustainability boundaries. *River Research and Applications*, **26**(8), 1052–1063.

Richter, Brian D, Warner, Andrew T, Meyer, Judy L, & Lutz, Kim. 2006. A collaborative and adaptive process for developing environmental flow recommendations. *River research and applications*, **22**(3), 297–318.

Richter, Brian D, Davis, MM, Apse, Colin, & Konrad, Christopher. 2012. A presumptive standard for environmental flow protection. *River Research and Applications*, **28**(8), 1312–1321.

RWP. 2011. *Middle Ocmulgee, Regional Water Plan*.

Sabo, John L, & Post, David M. 2008. Quantifying periodic, stochastic, and catastrophic environmental variation. *Ecological Monographs*, **78**(1), 19–40.

Schmid, Bernhard H, & Koskiaho, Jari. 2006. Artificial neural network modeling of dissolved oxygen in a wetland pond: the case of Hovi, Finland. *Journal of Hydrologic Engineering*, **11**(2), 188–192.

Seebacher, Frank, Guderley, Helga, Elsey, Ruth M, & Trosclair, Phillip L. 2003. Seasonal acclimatisation of muscle metabolic enzymes in a reptile (Alligator mississippiensis). *Journal of Experimental Biology*, **206**(7), 1193–1200.

Shafroth, Patrick B, Wilcox, Andrew C, Lytle, David A, Hickey, John T, Andersen, Douglas C, Beauchamp, Vanessa B, Hautzinger, Andrew, McMULLEN, LAURA E, & Warner, Andrew. 2010. Ecosystem effects of environmental flows: modelling and experimental floods in a dryland river. *Freshwater Biology*, **55**(1), 68–85.

Shim, Natalia V, & Tollner, Ernest W. 2014. Application of recurrent neural networks for predicting water quality constituents of collected runoff from windrow composting pad. *In: 21st Century Watershed Technology Conferenceand Workshop Improving Water Quality and the Environment Conference Proceedings, University of Waikato, New Zealand*. ASABE.

Singh, Kunwar P, Basant, Ankita, Malik, Amrita, & Jain, Gunja. 2009. Artificial neural network modeling of the river water qualityâĂŤa case study. *Ecological Modelling*, **220**(6), 888–895.

Singh, Kunwar P, Basant, Nikita, Malik, Amrita, & Jain, Gunja. 2010. Modeling the performance of âĂIJup-flow anaerobic sludge blanketâĂİ reactor based wastewater treatment plant using linear and nonlinear approachesâĂŤa case study. *Analytica chimica acta*, **658**(1), 1–11.

Smakhtin, Vladimir U. 2001. Low flow hydrology: a review. *Journal of hydrology*, **240**(3), 147–186.

Ssegane, H, Tollner, EW, McCutcheon, SC, *et al.* 2009. Riparian sediment delivery ratio: stiff diagrams and artificial neural networks. *Transactions of the ASABE*, **52**(6), 1885–1893.

Strayer, David L, & Dudgeon, David. 2010. Freshwater biodiversity conservation: recent progress and future challenges. *Journal of the North American Benthological Society*, **29**(1), 344–358.

Suen, Jian-Ping, & Eheart, J Wayland. 2003. Evaluation of neural networks for modeling nitrate concentrations in rivers. *Journal of water resources planning and management*, **129**(6), 505–510.

Swinson, K. 2014. Contour Generation to Support Calculation of Exposed Snail Habitat in the Coosa River, Alabama. *In: American Water Resources Association Conference,Tysons Corner, VA, 3-6 November 2014*.

Tharme, Rebecca E. 2003. A global perspective on environmental flow assessment: emerging trends in the development and application of environmental flow methodologies for rivers. *River research and applications*, **19**(5-6), 397–441.

Tollner, Ernest W, & Das, Keshav C. 2004. Predicting runoff from a yard waste windrow composting pad. *Transactions of the ASAE*, **47**(6), 1953–1961.

Toth, E, Brath, A, & Montanari, A. 2000. Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology*, **239**(1), 132–147.

Travnichek, Vincent H, & Maceina, Michael J. 1994. Comparison of flow regulation effects on fish assemblages in shallow and deep water habitats in the Tallapoosa River, Alabama. *Journal of Freshwater Ecology*, **9**(3), 207–216.

UORWPC. 2011. *Upper Oconee Regional Water Plan*.

Verma, Anoop, Wei, Xiupeng, & Kusiak, Andrew. 2013. Predicting the total suspended solids in wastewater: a data-mining approach. *Engineering Applications of Artificial Intelligence*, **26**(4), 1366–1372.

Vogel, Richard M, Sieber, Jack, Archfield, Stacey A, Smith, Mark P, Apse, Colin D, & Huber-Lee, Annette. 2007. Relations among storage, yield, and instream flow. *Water Resources Research*, **43**(5).

Wheatcroft, Robert A, Hatten, Jeff A, Pasternack, Gregory B, Warrick, Jonathan A, *et al.* 2010. The role of effective discharge in the ocean delivery of particulate organic carbon by small, mountainous river systems. *Limnology and Oceanography*, **55**(1), 161–171.

Wilson, Bruce G, Haralampides, Katy, & Levesque, Sheila. 2004. Stormwater runoff from open windrow composting facilities. *Journal of Environmental Engineering and Science*, **3**(6), 537–540.

Wolman, M Gordon, & Miller, John P. 1960. Magnitude and frequency of forces in geomorphic processes. *The Journal of Geology*, **68**(1), 54–74.

Xie, Hang, Tang, Hao, & Liao, Yu-He. 2009. Time series prediction based on NARX neural networks: An advanced approach. *Pages 1275–1279 of: Machine Learning and Cybernetics, 2009 International Conference on*, vol. 3. IEEE.

Yao, H, & Georgakakos, A. 2001. Assessment of Folsom Lake response to historical and potential future climate scenarios: 2. Reservoir management. *Journal of Hydrology*, **249**(1), 176–196.

# Appendices

# Appendix A

# Supplementary Time Series Data

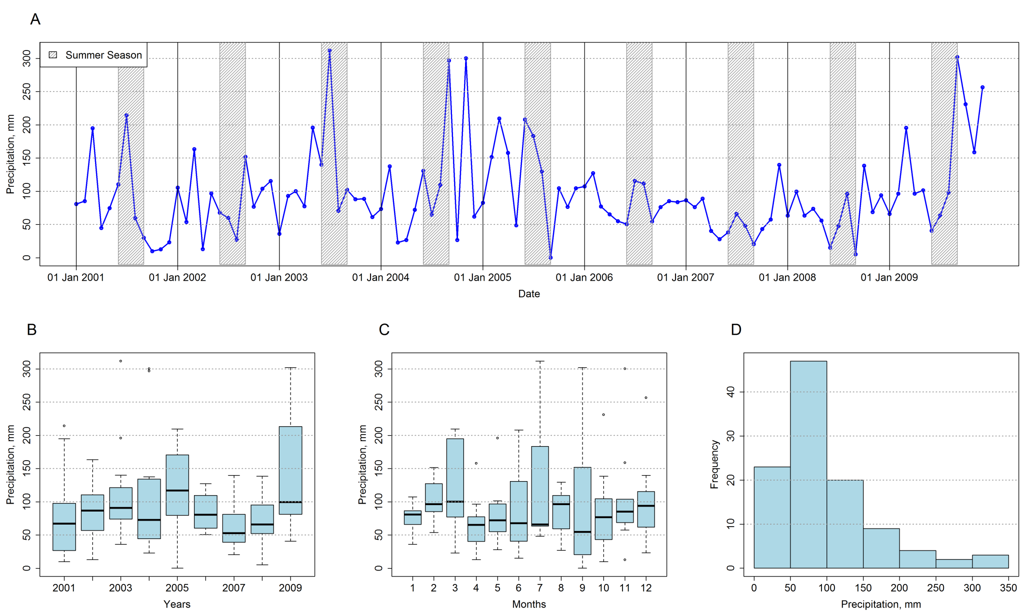



Figure A1: Time Series Data of precipitation (mm).

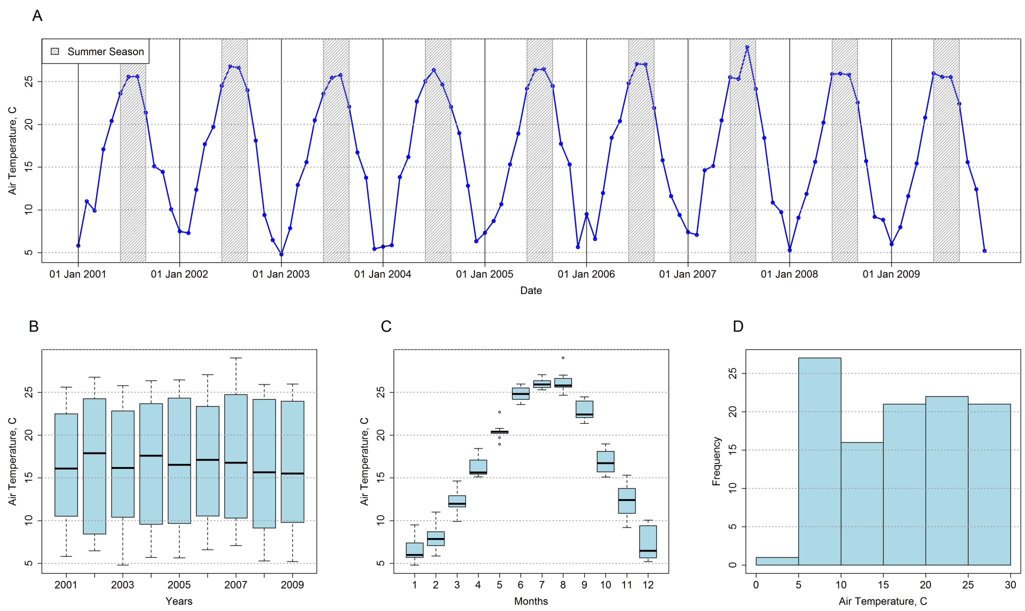



Figure A2: Time Series Data of air temperature (ÂřC).

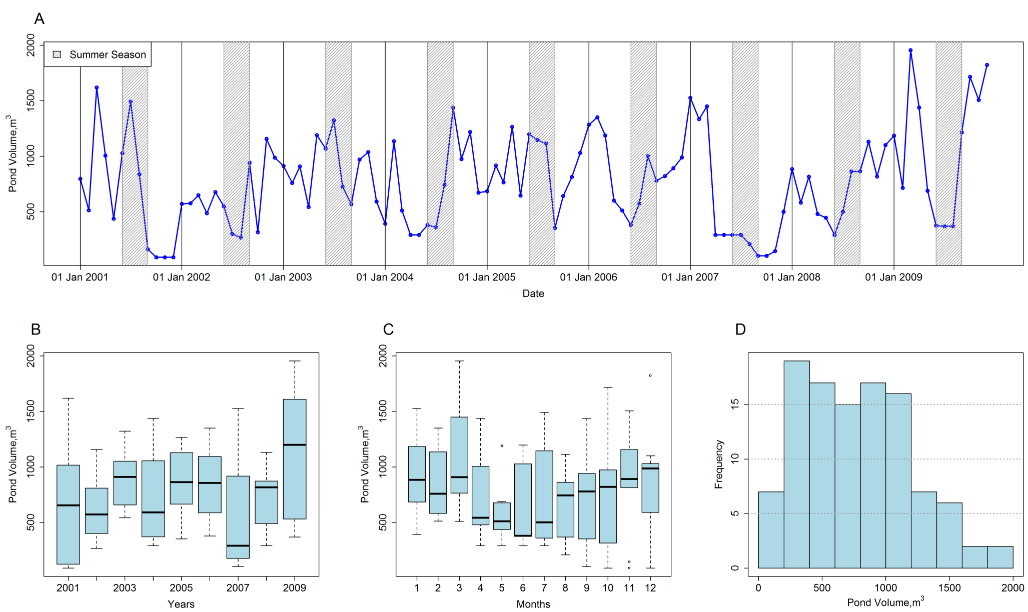



Figure A3: Time Series Data of pond volume ($m^3$).

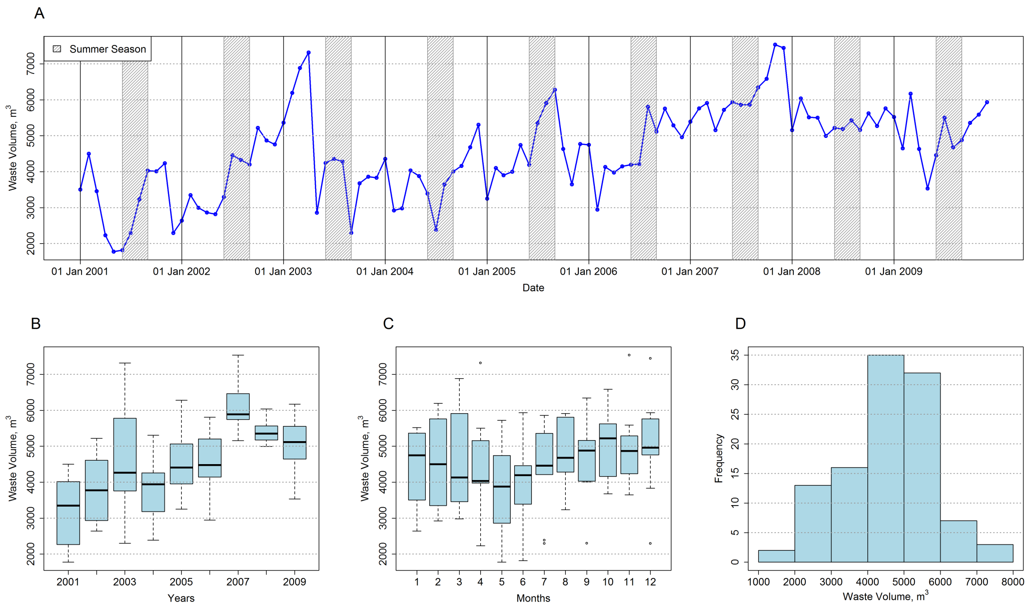



Figure A4: Time Series Data of waste volume ($m^3$).

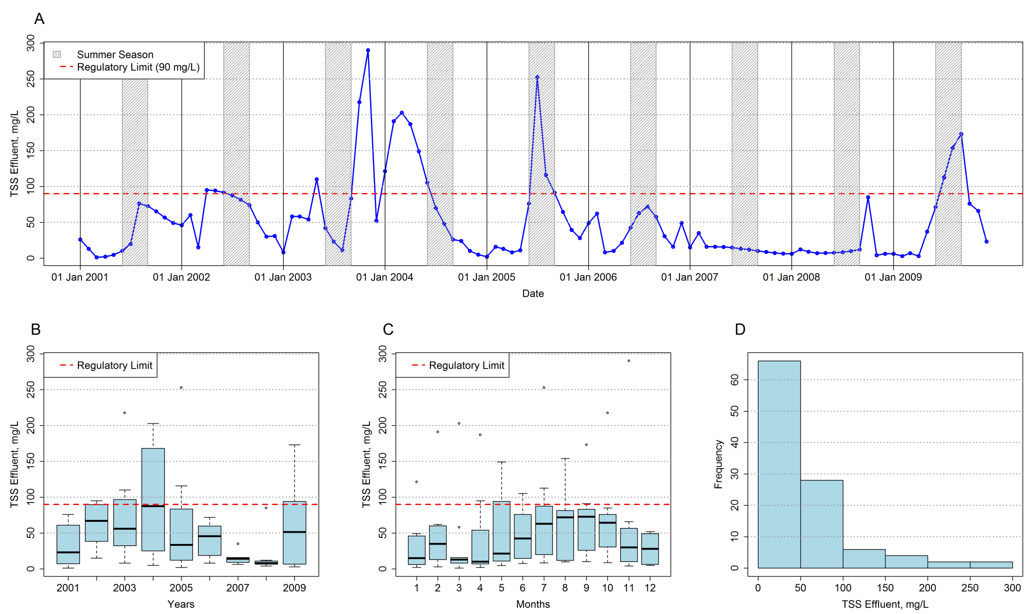



Figure A5: Time Series Data of Total Suspended Solids (mg/L).

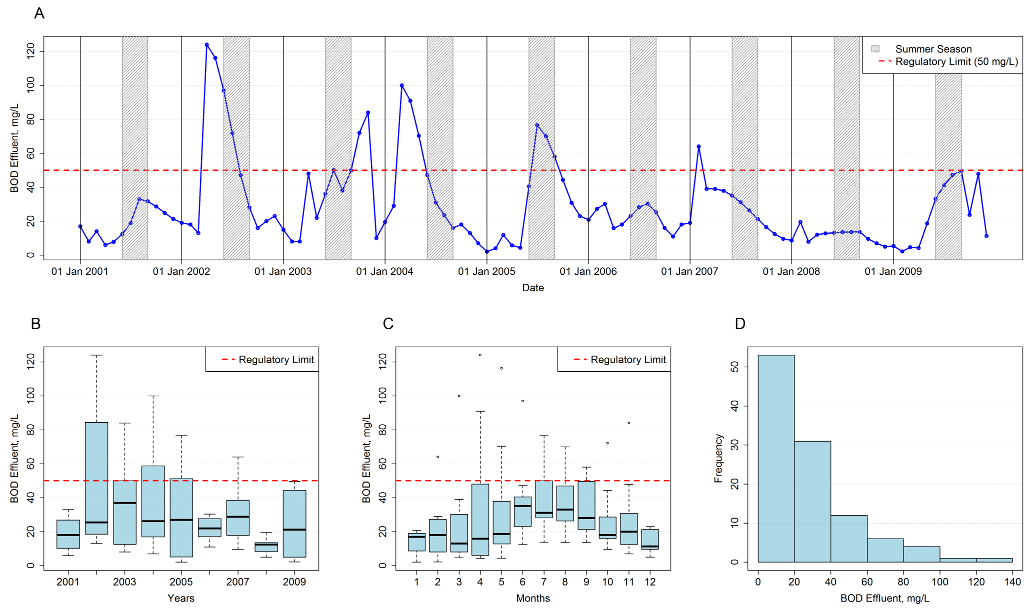



Figure A6: Time Series Data of Biological Oxygen Demand (mg/L).

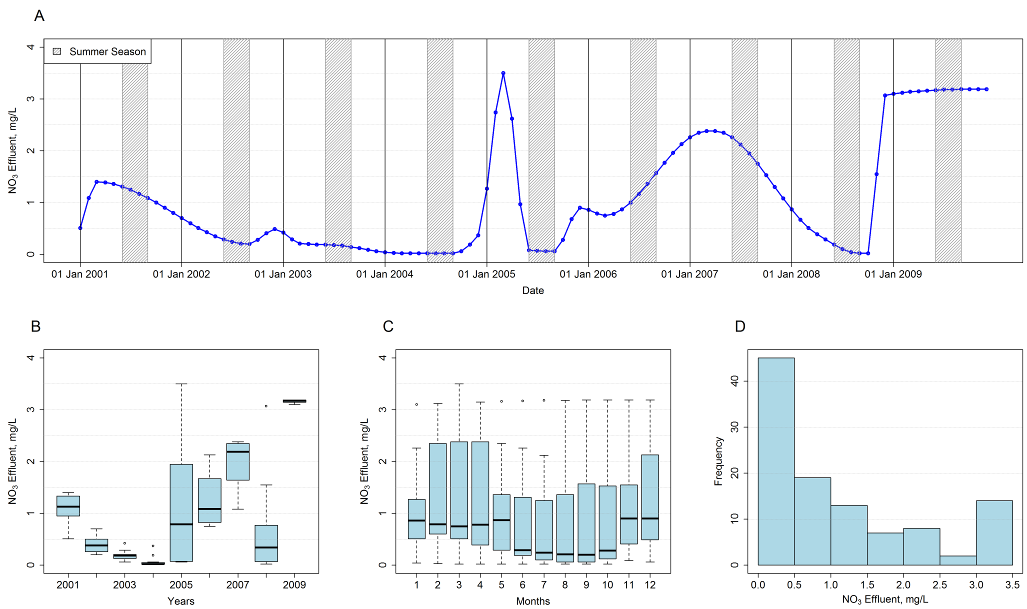
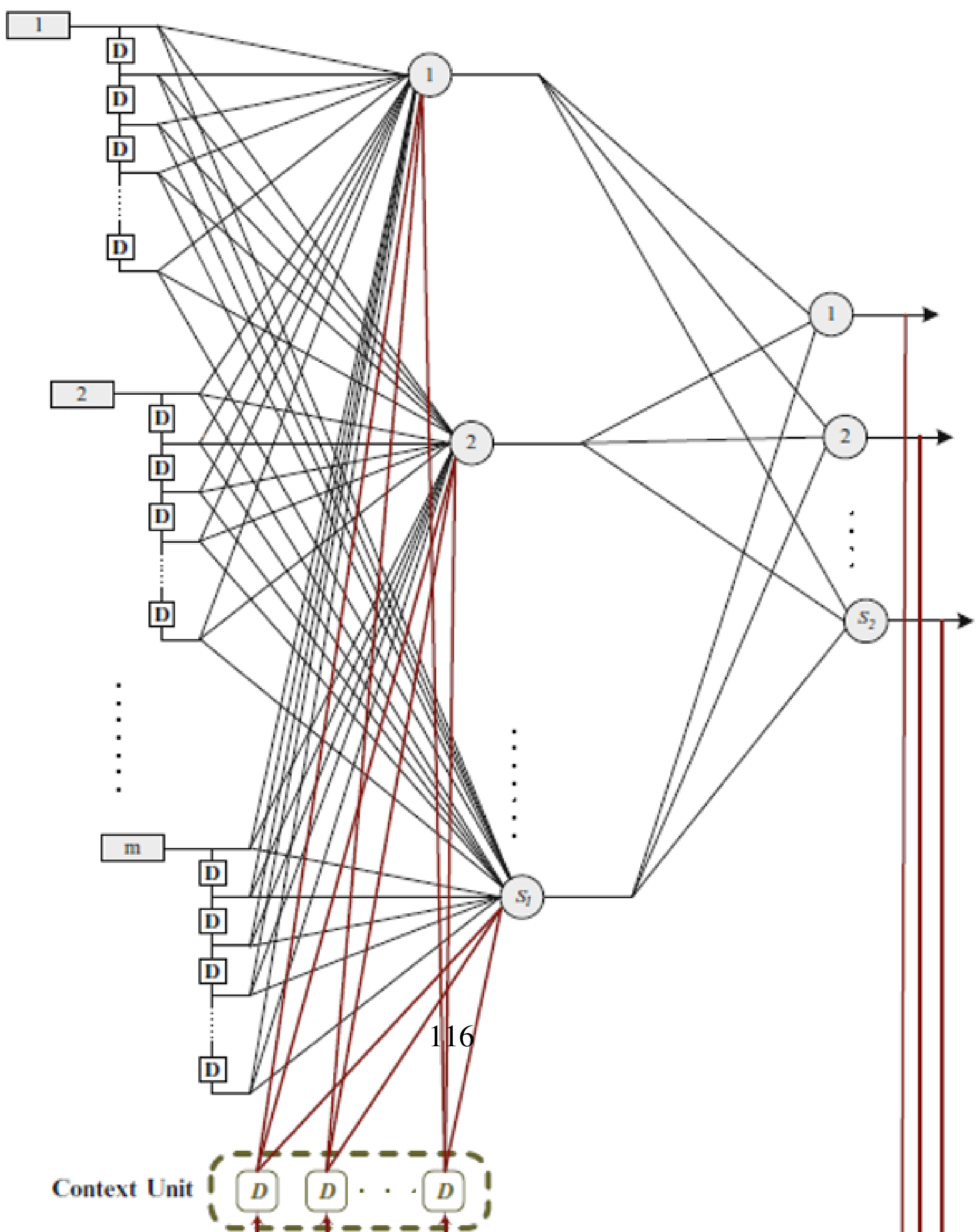



Figure A7: Time Series Data of Nitrate (mg/L).

# Appendix B

# Supplementary Figures



**Input-Hidden Connection Weights**

| X \ Y | Hidden A | Hidden B | Hidden C | Hidden D | Hidden E |
|---|---|---|---|---|---|
| **Input 1** | -0.93 | -1.49 | 0.37 | -0.91 | 0.37 |
| **Input 2** | -0.57 | -1.96 | -0.14 | 1.18 | 1.26 |
| **Input 3** | -0.85 | 1.74 | -1.86 | -0.05 | 0.10 |
| **Input 4** | 0.25 | -3.01 | -0.99 | -1.34 | -1.65 |
| **Input 5** | -0.82 | 0.09 | 0.86 | -0.41 | -0.05 |

X

**Hidden-Output Connection Weights**

| | Hidden A | Hidden B | Hidden C | Hidden D | Hidden E |
|---|---|---|---|---|---|
| **Output** | -1.75 | -1.08 | -1.13 | -2.90 | 3.37 |

**Connection Weight Products**

| X \ Y | Hidden A | Hidden B | Hidden C | Hidden D | Hidden E |
|---|---|---|---|---|---|
| **Input 1** | 1.63 | 1.62 | -0.42 | 2.64 | 1.24 |
| **Input 2** | 1.00 | 2.12 | 0.16 | -3.43 | 4.25 |
| **Input 3** | 1.48 | -1.89 | 2.10 | 0.14 | 0.34 |
| **Input 4** | -0.43 | 3.26 | 1.12 | 3.90 | -5.57 |
| **Input 5** | 1.43 | -0.09 | -0.97 | 1.18 | -0.16 |

$$Input_X = \sum_{Y=A}^{E} Hidden_{XY}$$

$$Input_X = \sum_{Y=A}^{E} \frac{|Hidden_{XY}|}{\sum_{Z=1}^{5} |Hidden_{ZY}|}$$

| | Importance | Rank |
|---|---|---|
| **Input 1** | 6.71 | 1 |
| **Input 2** | 4.10 | 2 |
| **Input 3** | 2.18 | 3 |
| **Input 4** | 2.28 | 4 |
| **Input 5** | 1.38 | 5 |

**Connection Weight Approach**

| | Importance | Rank |
|---|---|---|
| **Input 1** | 0.88 | 4 |
| **Input 2** | 1.11 | 2 |
| **Input 3** | 0.94 | 3 |
| **Input 4** | 1.50 | 1 |
| **Input 5** | 0.57 | 5 |

**Garson's Algorithm**

Figure B2: Calculation of weight importance in Olden's algorithm (Connection Weight Approach) and Garson's algorithm (Olden *et al.*, 2004)

.

# Appendix C

# R Script

```
###############################################################
#IMPORT DATA
###############################################################
rm(list=ls(all=TRUE)) #Clear memory

################################################################
#Set up working directory
################################################################

setwd("___________")

library(MASS)

#################################################################
#Habitat rating curve (Q in cfs, area in acres)
```

```
##################################################################
dummy <- read.csv("HabitatResults_2016-10-27_Import.csv", header=TRUE, dec=".")
habitat.in <- data.matrix(dummy)

##################################################################
#Discharge data
##################################################################
#Import daily discharge data in cfs
dummy2 <- read.csv("MIDO_1938-1997.csv", header=FALSE, dec=".")
Q <- data.matrix(dummy2) #Convert to matrix format
Q_cms <- Q*0.3048^3 #Convert to cubic meters per sec
n <- length(Q) #Number of data points
years <- n/365 #Number of years in data set
Qc <- c() #Compile into a single time series
for(i in 1:years){Qc <- c(Qc,Q[,i])}
Qc_cms <- Qc*0.3048^3

##################################################################
#Temporal Properties
##################################################################
t0 <- 1938 #First year in record
tf <- t0 + years - 1  #Last year in record
month.day <- c(1,32,60,91,121,152,182,213,244,274,305,335) #First day of each month
days <- c(31,28,31,30,31,30,31,31,30,31,30,31) #Days per month
month.end <- month.day + days - 1 #Last day of each month
month.label <- c("J","F","M","A","M","J","J","A","S","O","N","D")
```

```
#####################################################################################
#INITIAL DISCHARGE STATISTICS
#####################################################################################
#Basic Discharge Properties
Qmean <- mean(Qc); Qmean_cms <- Qmean*0.3048^3
Qmin <- min(Qc); Qmin_cms <- Qmin*0.3048^3
Qmax <- max(Qc); Qmax_cms <- Qmax*0.3048^3
Qmedian <- median(Qc); Qmedian_cms <- Qmedian*0.3048^3
summary(Qc)
summary(Qc_cms)

#   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#    8.2   219.0   350.0   521.3   562.0 12600.0

#Compute monthly discharge statistics
Qm.mean <- c(); Qm.median <- c(); Qm.min <- c(); Qm.max <- c();
for(i in 1:12){
Qm.mean[i] <- mean(apply(Q[month.day[i]:month.end[i],],2,mean))
Qm.median[i] <- median(apply(Q[month.day[i]:month.end[i],],2,mean))
Qm.min[i] <- min(apply(Q[month.day[i]:month.end[i],],2,mean))
Qm.max[i] <- max(apply(Q[month.day[i]:month.end[i],],2,mean))
}

Qm.mean_cms <- Qm.mean*0.3048^3
Qm.median_cms <- Qm.median*0.3048^3
```

```
Qm.min_cms <- Qm.min*0.3048^3
Qm.max_cms <- Qm.max*0.3048^3

################################################################
#FLOW MANAGEMENT
###############################################################
#Specify the general withdrawal parameters for flow modification
Qpump.MGD <- 60 #Pump capacity (MGD)
Qpump <- Qpump.MGD*(10^6)/(86400*7.48) #Pump capacity (cfs)
Qpump_cms <- Qpump*0.3048^3  #Pump capacity (cms)
Qpump.max <- 500 #Maximum river discharge for withdrawal (cfs)
Monthly7Q10 <- c(rep(247, 31), rep(283, 28), rep(316, 31),
            rep(289,30), rep(185,31), rep(133,30), rep(113,31),
rep(67,31), rep(53,30), rep(88,31), rep(146,30), rep(175,31))
Monthly7Q10_cms <- Monthly7Q10*0.3048^3
Annual7Q10 <- 37
Annual7Q10_cms <- Annual7Q10*0.3048^3

##################################################################
# Calculate how many days the flow was below annual 7Q10
##################################################################
Qbelow7Q10 <- Qc[which(Qc<Annual7Q10)]
Qbelow7Q10sort <- sort(Qbelow7Q10)
k = length(Qbelow7Q10)
m = 1 + 3.3*log10(k) # number of breaks for the histogram
```

```
##################################################################
#FIGURE - HISTOGRAM OF LOW FLOWS BELOW ANNUAL 7Q10
##################################################################
dev.new()
h = hist(Qbelow7Q10sort,breaks=m,ylab="Frequency",xlab="Flow, cfs",
        col="grey",main="")
for(h in seq(10,40,10)){
abline(h=h, col='gray60', lwd=1.5, lty=3)
}
box(col = 1)

##################################################################
#Parameterize eflow scenarios
##################################################################
nMFL <- 100+1
AMF <- seq(0, 1000, length.out=nMFL)
MMF <- matrix(0,nrow=12,ncol=nMFL)
for(i in 1:nMFL){MMF[,i] <- Qm.min + (i-1)/(nMFL-1) * (Qm.max-Qm.min)}
SB <- seq(0.50, 0, length.out=nMFL)

##################################################################
#Create empty matrices to store results
#Empty matrix to store withdrawal discharge
Qw.AMF <- array(0, dim=c(365, years, nMFL));
Qw.MMF <- array(0, dim=c(365, years, nMFL));
Qw.SB <- array(0, dim=c(365, years, nMFL))
```

```
Qw.AMFc <- array(0, dim=c(365, years, nMFL));
Qw.MMFc <- array(0, dim=c(365, years, nMFL));
Qw.SBc <- array(0, dim=c(365, years, nMFL))

#Empty matrix to store river discharge
Qr.AMF <- array(0, dim=c(365, years, nMFL));
Qr.MMF <- array(0, dim=c(365, years, nMFL));
Qr.SB <- array(0, dim=c(365, years, nMFL))
Qr.AMFc <- array(0, dim=c(365, years, nMFL));
Qr.MMFc <- array(0, dim=c(365, years, nMFL));
Qr.SBc <- array(0, dim=c(365, years, nMFL))

#################################################################
#Functions for computing flow management - UNCONSTRAINED
#Annual minimum flows
flow.AMF <- function(Q, MFL, Qpump){ifelse(Q>(MFL+Qpump), Qpump,
                ifelse(Q>MFL,Q-MFL,0))}
#Monthly minimum flows
flow.MMF <- function(Q, MMF, Qpump, month.day, month.end){
Qw.temp <- matrix(0, nrow=365, ncol=years)
for(i in 1:365){
m <- which(month.day<=i & month.end>=i)
Qw.temp[i,] <- ifelse(Q[i,]>(MMF[m]+Qpump), Qpump,
            ifelse(Q[i,]>MMF[m],Q[i,]-MMF[m],0))
}
Qw.temp
```

```
}
#Sustainability boundaries
flow.SB <- function(Q, SB, Qpump){ifelse(SB*Q > Qpump, Qpump, SB*Q)}

################################################################
#Loop over each flow management scenario to modify hydrographs
for (k in 1:nMFL){
#Compute withdrawal rates
Qw.AMF[,,k] <- flow.AMF(Q, AMF[k], Qpump)
Qw.MMF[,,k] <- flow.MMF(Q, MMF[,k], Qpump, month.day, month.end)
Qw.SB[,,k] <- flow.SB(Q, SB[k], Qpump)

#Compute river discharge associated with withdrawal
Qr.AMF[,,k] <- Q - Qw.AMF[,,k]
Qr.MMF[,,k] <- Q - Qw.MMF[,,k]
Qr.SB[,,k] <- Q - Qw.SB[,,k]
}

#Compute AVERAGE withdrawal rate (MGD)
AMF.rate <- apply(Qw.AMF, 3, mean) * 86400 * 7.48 / (10^6)
MMF.rate <- apply(Qw.MMF, 3, mean) * 86400 * 7.48 / (10^6)
SB.rate <- apply(Qw.SB, 3, mean) * 86400 * 7.48 / (10^6)

#Compute withdrawal rate BY YEAR (MGD)
AMF.rate2 <- apply(Qw.AMF, 2:3, mean) * 86400 * 7.48 / (10^6)
MMF.rate2 <- apply(Qw.MMF, 2:3, mean) * 86400 * 7.48 / (10^6)
```

```
SB.rate2 <- apply(Qw.SB, 2:3, mean) * 86400 * 7.48 / (10^6)

#Compute MINIMUM withdrawal rate (MGD)
AMF.rate3 <- apply(AMF.rate2, 2, min)
MMF.rate3 <- apply(MMF.rate2, 2, min)
SB.rate3 <- apply(SB.rate2, 2, min)

#Compute AVERAGE river discharge (cfs)
Qr.AMF.avg <- apply(Qr.AMF, 3, mean)
Qr.MMF.avg <- apply(Qr.MMF, 3, mean)
Qr.SB.avg <- apply(Qr.SB, 3, mean)

#Compute MEDIAN river discharge (cfs)
Qr.AMF.med <- apply(Qr.AMF, 3, median)
Qr.MMF.med <- apply(Qr.MMF, 3, median)
Qr.SB.med <- apply(Qr.SB, 3, median)

######################################################################
#FUNCTION FOR HABITAT PROVISION
#####################################################################
#Set parameters for linear interpolation of habitat
Qh <- habitat.in[,1]
nhab <- length(Qh)
habitat.temp <- habitat.in[,2:ncol(habitat.in)]
habitat.slope <- (habitat.temp[-1,] - habitat.temp[-nhab,]) / (Qh[-1] - Qh[-nhab])
habitat.int <- habitat.temp[-1,] - habitat.slope * Qh[-1]
```

```
##################################################################
#Function for computing habitat at any discharge
#Q is any value of discharge in cfs
#Outputs are a vector of habitat quantities in acres at that discharge in cfs
#Five values are output: total, shallow-fast, deep-fast, shallow-slow, other
habitat <- function(Q){
loc <- max(which(Qh <= Q))
habitat.slope[loc,] * Q + habitat.int[loc,]
}

#Test function
#habitat(Qc[3])

#####################################################################
#EFFECTIVENESS ANALYSIS FOR HABITAT
####################################################################
#Define bins for effectiveness analysis
#Qbins <- Qh
#kernel.total <- density(Qc, kernel="gaussian", from=0, to=Qmax)
#Qbins <- kernel.total$x
# Qbins <- c(seq(0,9,1), seq(10,500,10), seq(500,900,100),
        seq(1000,9000,1000), seq(10000,18000,1000))
Qbins <- c(seq(0,9,1), seq(10,500,10), seq(600,900,100),
        seq(1000,19000,1000))
Qbins.mid <- hist(Q, Qbins, plot = FALSE)$mids
```

```
nbins <- length(Qbins)

####################################################################
#Compute UNALTERED values of the area under the effectiveness curve
f.total <- (hist(Q, Qbins, plot = FALSE)$counts) / n
habitat.rating <- matrix(0, nrow=nbins-1, ncol=ncol(habitat.in)-1)
E.total <- matrix(0, nrow=nbins-1, ncol=ncol(habitat.in)-1)
for(i in 1:(nbins-1)){habitat.rating[i,] <- habitat(Qbins.mid[i])}
for(i in 1:(ncol(habitat.in)-1)){E.total[,i] <- f.total * habitat.rating[,i]}
Aeff.hab.un <- apply(E.total, 2, sum)

####################################################################
#Compute effective habitat metric for all flow scenarios
Aeff.hab <- array(0, dim=c(nMFL,3,ncol(habitat.in)-1))
for(i in 1:(ncol(habitat.in)-1)){
for(j in 1:nMFL){
#Frequency distributions
ftemp1 <- (hist(Qr.AMF[,,j], Qbins, plot = FALSE)$counts) / n
ftemp2 <- (hist(Qr.MMF[,,j], Qbins, plot = FALSE)$counts) / n
ftemp3 <- (hist(Qr.SB[,,j], Qbins, plot = FALSE)$counts) / n

#Effectiveness curves
Etemp1 <- ftemp1 * habitat.rating[,i]
Etemp2 <- ftemp2 * habitat.rating[,i]
Etemp3 <- ftemp3 * habitat.rating[,i]
```

```
#Effectiveness metrics
Aeff.hab[j,1,i] <- sum(Etemp1)
Aeff.hab[j,2,i] <- sum(Etemp2)
Aeff.hab[j,3,i] <- sum(Etemp3)
}
}

#Normalize effectiveness metrics for flow scenarios
Aeff.hab.norm <- array(0, dim=c(nMFL,3,ncol(habitat.in)-1))
for(i in 1:(ncol(habitat.in)-1)){Aeff.hab.norm[,,i] <-
(Aeff.hab.un[i] - abs(Aeff.hab[,,i] - Aeff.hab.un[i])) / Aeff.hab.un[i]}

##################################################################
#Deterministic view of habitat analysis
hab.avg <- array(0, dim=c(nMFL,3,ncol(habitat.in)-1))
hab.med <- array(0, dim=c(nMFL,3,ncol(habitat.in)-1))
for(j in 1:nMFL){
#Average habitat
hab.avg[j,1,] <- habitat(Qr.AMF.avg[j])
hab.avg[j,2,] <- habitat(Qr.MMF.avg[j])
hab.avg[j,3,] <- habitat(Qr.SB.avg[j])

#Median habitat
hab.med[j,1,] <- habitat(Qr.AMF.med[j])
hab.med[j,2,] <- habitat(Qr.MMF.med[j])
hab.med[j,3,] <- habitat(Qr.SB.med[j])
```

```
}

#Unaltered deterministic habitat
hab.avg.un <- habitat(Qmean)
hab.med.un <- habitat(Qmedian)

#PLOTTING
##########################################################
#Specify withdrawal rates for presenting trade-offs
#Qw.all <- cbind(AMF.rate, MMF.rate, SB.rate)
Qw.all <- cbind(AMF.rate3, MMF.rate3, SB.rate3)

###########
#General plotting options
#col.plot <- c("darkred", "darkgreen", "darkblue")
col.plot <- c("goldenrod", "darkorchid2", "darkblue")
#xlab1 <- c("Average Annual Withdrawal Rate")
xlab1 <- c("Minimum Annual Withdrawal Rate")
ylab1 <- c("Condition Index")
labels.regime <- c("Annual Minimum",
                "Monthly Minimum", "Percent of Flow")

##########################################################
#FIGURE 1 - ENVELOPE HYDROGRAPH in cms
##########################################################
win.graph(14,7)
```

```
par(mar=c(4,4,1,1), cex=1.5)

plot(c(1,1), c(1,1), type="n", log="y", lwd=3, xlim=c(1,365),
    ylim=c(0.2,400), axes=FALSE, xlab="", ylab="Discharge (L/s)", main="")
axis(1, at=month.day, labels=month.label, tck=0)
axis(2)
box()
abline(v=month.day, col = "lightgray", lty=3)

xtemp <- c(seq(1,365), seq(365,1))
ytemp.min <- apply(Q_cms, 1, min)
ytemp.max <- apply(Q_cms, 1, max)
ytemp.range <- c(ytemp.min, rev(ytemp.max))
polygon(xtemp, ytemp.range, col="grey", border=NA)
lines(seq(1,365),apply(Q_cms, 1, median), lwd=3, col=1)
abline(h=Qmean_cms, col="black", lwd=3, lty=3)
abline(h=Qpump_cms, col="red", lwd=3, lty=3)
abline(h=Annual7Q10_cms, col = "blue", lwd=3, lty=3)
lines(seq(1,365),Monthly7Q10_cms, col="forestgreen", lwd=3)
legend("bottomleft", legend=c("Range", "Median Daily Discharge",
        "Long-term Average", "60 MGD", "Annual 7Q10", "Monthly 7Q10"),
fill=c("grey", "black", NA, NA, NA, NA),
        border=c("black","black",NA,NA,NA,NA),
lty=c(NA,NA,3,3,3,1), lwd=c(NA,NA,3,3,3,3),
        col=c(NA,NA,"black", "red","blue","forestgreen"),
        bg="white", cex=0.7)
```

# Appendix D

# Python Script

```
#------Inputs---------#

indir = r"_____"  # directory containing the input "dp" and "vp" rasters

minNum = 1  # minimum integer value from the raster names to process

maxNum = 79 # maximum integer value from the raster names to process

outdir = r"__________"    #where to save the output rasters

outtable = r"_________"   #where you want to save the output table containing

                          #the percentages


#-----Processing-------#


#----Import Arcpy, Set Environment Variables-----#


print "importing arcpy..."

import arcpy

import os
```

```
arcpy.CheckOutExtension("spatial")
from arcpy.sa import *
from arcpy import env
arcpy.env.workspace = "in_memory"
env.scratchWorkspace = "in_memory" #Do all processing in memory
env.overwriteOutput = True

#-------Define some useful functions-------#
def scratch(x):
    """Creates Scratch Filenames in RAM"""
    return arcpy.CreateScratchName(x,'','',"in_memory")

def vals(fc,field):
    """Extracts field values from arcgis tables"""
    values = [row[0] for row in arcpy.da.SearchCursor(fc, (field))]
    return values

#-------Get list all all numbers from min to max------#
allnum = [format(x,'03') for x in range(minNum,maxNum+1)]

#-------Make empty table to record percentages---------#
temptab = arcpy.CreateTable_management("in_memory",'temptab')
#make a temporary table (stored in memory)
arcpy.AddField_management(temptab,'simnum','TEXT',field_length = 10)
#add field to store the simulation number
arcpy.AddField_management(temptab,'sum','FLOAT')
```

```
arcpy.AddField_management(temptab,'sf_sum','FLOAT')
arcpy.AddField_management(temptab,'df_sum','FLOAT')
arcpy.AddField_management(temptab,'ss_sum','FLOAT')
arcpy.AddField_management(temptab,'sf_per','FLOAT')
arcpy.AddField_management(temptab,'df_per','FLOAT')
arcpy.AddField_management(temptab,'ss_perc','FLOAT')

#-----Loop through each number (turning it into the input filenames)-----#
for n in allnum:
    print "working on number %s" % (n,)
    v_PF = os.path.join(indir,"vp%s" % (n))
    d_PF = os.path.join(indir,"dp%s" % (n))

    # Process: Raster Calculator
    Q10_v_x1000 = Raster(v_PF) * 1000

    # Process: Raster Calculator (2)
    Q10_d_x1000 = Raster(d_PF) * 1000

########################
# Shallow fast habitat #
########################

    # Process: Raster Calculator (3)
    Q10_v_x1000_shallowfast_true = Q10_v_x1000 >= 1800
    Q10_v_x1000_shallowfast_true)
```

```
    # Process: Raster Calculator (4)
    Q10_d_x1000_shallowfast_true = Q10_d_x1000 <= 1150
    Q10_d_x1000_shallowfast_true)

    # Process: Raster Calculator (5)
    Q10_shallowfast = (Q10_v_x1000_shallowfast_true == 1) &
                      (Q10_d_x1000_shallowfast_true == 1)

#######################
# Deep fast habitat   #
#######################

    # Process: Raster Calculator (6)
    Q10_d_x1000_deepfast_true = Q10_d_x1000 >= 1150

    # Process: Raster Calculator (7)
    Q10_v_x1000_deepfast_true = Q10_v_x1000 >  1480

    # Process: Raster Calculator (8)
    Q10_deepfast = (Q10_d_x1000_deepfast_true == 1) &
                   (Q10_v_x1000_deepfast_true == 1)

#################
# Shallow Slow  #
#################
```

```
# Process: Raster Calculator (9)
Q10_d_x1000_shallowslow_true = Q10_d_x1000 < 1150

# Process: Raster Calculator (10)
Q10_v_x1000_shallowslow_true = Q10_v_x1000 < 1150

# Process: Raster Calculator (11)
Q10_shallowslow = (Q10_d_x1000_shallowslow_true == 1) &
                  (Q10_v_x1000_shallowslow_true == 1)

############    Save Rasters    #################

Q10_shallowslow.save(os.path.join(outdir,"Q10_shallowslow_%s.tif" % n))
Q10_shallowfast.save(os.path.join(outdir,"Q10_shallowfast_%s.tif" % n))
Q10_deepfast.save(os.path.join(outdir,"Q10_deepfast_%s.tif" % n))

############   Extract the percentages  #############

shallfast = arcpy.CopyRaster_management(Q10_shallowfast,scratch('shallfast'))
sf_count = vals(shallfast,'Count')
if sf_count[0]==sum(sf_count):
    sf_count_true = 0
else:
    sf_count_true = sf_count[1]
sf_perc_val = sf_count_true/sum(sf_count)
```

```
    deepfast = arcpy.CopyRaster_management(Q10_deepfast,scratch('deepfast'))
    df_count = vals(deepfast,'Count')
    if df_count[0]==sum(df_count):
        df_count_true = 0
    else:
        df_count_true = df_count[1]
    df_perc_val = df_count_true/sum(df_count)

    shallslow = arcpy.CopyRaster_management(Q10_shallowslow,scratch('shallslow'))
    ss_count = vals(shallslow,'Count')
    if ss_count[0] == sum(ss_count):
        ss_count_true = 0
    else:
        ss_count_true = ss_count[1]
    ss_perc_val = ss_count_true/sum(ss_count)

    ############  Append values to temporary table  ##########
    ############  Add values to table               ##########

    with arcpy.da.InsertCursor(temptab,['simnum', 'sum', 'sf_sum', 'df_sum',
                              'ss_sum', 'sf_per', 'df_per', 'ss_perc']) as icurs:
    icurs.insertRow([n,sum(ss_count), sf_count_true, df_count_true, ss_count_true,
                sf_perc_val, df_perc_val,  ss_perc_val])

#####  Copy temperary table to output location ##########
```

```
arcpy.CopyRows_management(temptab,outtable)
arcpy.Delete_management("in_memory")
print "done!"
```