

RAPID FLOOD DAMAGE PREDICTION IN SHORT LEAD-TIME SCENARIOS

by

JOSEPH LEE GUTENSON

ANDREW N. S. ERNEST, COMMITTEE CHAIR

ANDREW J. GRAETTINGER

GARY P. MOYNIHAN

GLENN A. TOOTLE

DAVID R. MAIDMENT

LINDELL E. ORMSBEE

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Civil, Construction, and Environmental Engineering
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2016

Copyright Joseph Lee Gutenson 2016
ALL RIGHTS RESERVED

ABSTRACT

Research on short lead-time flood inundation and damage assessment traditionally focuses on developing tools for long term planning. Studies that investigate the ability to provide short lead-time analyses are uncommon because of sparcity and spatial inconsistency of short-term hydrologic forecasts. Recent advances in continental scale hydrology make possible the ability to predict discharge at nearly all locations in the conterminous United States (CONUS). Theoretically, these discharge outputs enable investigation into the plausibility of short lead-time hydraulic, and damage analyses during flood events. A system predicting the hydraulics and damage potential of floods is only feasible after addressing a number of limitations. For instance, hydraulic models require a characterization of the stream channel, either through site surveys or through approximations. Further, the damage estimation methodology requires building inventories, which are developable by similar mechanisms. This dissertation advances knowledge on how to possibly address these limitations.

The first study in this dissertation investigates the use of hydraulic geometries. Hydraulic geometries relate river channel geometry to bankfull discharge. However, the research presented here indicates that hydraulic geometries may estimate both the channel geometry and multiple depths of flow, under certain geomorphic and anthropogenic constraints. Accurate channel geometries are necessary for hydraulic modeling. Depth of flow estimates are useful in developing stage-discharge rating curves and possibly as a standalone means of estimating inundation grids.

The second and third studies in this research look to investigate a framework for rapid flood damage assessment using public domain cadastral or parcel geospatial data. The second study discusses a fuzzy logic framework for meshing emergency response Address Points with parcel data and appropriate depth-damage relationships to determine both percent and fiscal impact of flood damage. Results of the second study highlight the effectiveness of determining flood damage with cadastral data using fuzzy text matching. The third study investigates what inputs from the cadastral data are necessary for the fuzzy logic framework to approximate a detailed flood damage investigation. Results indicate that fuzzy logic can approximate a detailed study when provided discrete use descriptions, market value, and square footage.

DEDICATION

I dedicate this research to my late grandfather, George A. Thompson, Sr.

LIST OF ABBREVIATIONS AND SYMBOLS

ADECA – Alabama Department of Economic and Community Affairs

AI – Artificial Intelligence

API – Applied Programming Interface

CONUS – Coterminous United States

DEM – Digital Elevation Model

EROM – Enhanced Runoff Method

FEMA – Federal Emergency Management Agency

FRR – Flood Risk Report

Hazus – United States Hazard Multi-Hazard

HEC-FDA – Hydrologic Engineering Center Flood Damage Analysis

HEC-FIA – Hydrologic Engineering Center Flood Impact Analysis

HEC-HMS – Hydrologic Engineering Center Hydrologic Modeling System

HEC-RAS – Hydrologic Engineering Center River Analysis System

HRRR – High Resolution Rapid Refresh

HSPF – Hydrologic Simulation Program FORTRAN

IDW – Inverse Distance Weighted

LDA – Latent Diriclet Allocation

NHDPlus Version 2 – National Hydrography Dataset Version 2

NFIE – National Flood Interoperability Experiment

OWDI – Open Water Data Initiative

PDA – Preliminary Damage Assessment

SPRNT – Simulation Program for River Networks

SWAT – Soil and Water Assessment Tool

USACE – United States Army Corps of Engineers

USGS – United States Geological Survey

VGI – Voluntary Geographic Information

VIC – Variable Infiltration Capacity Model

ACKNOWLEDGMENTS

I first want to acknowledge my advisor, Andy Ernest, for his continual support of me over the better part of seven years. I am proud to call him my colleague, mentor, and friend. I am also indebted to mother, Janet Haste, who raised me to be a lifelong learner and to work hard in whatever I pursue. I would like to thank my uncles George A. Thompson, Jr. and L. Mitchell Thompson for their words of encouragement over the years. I would also like to thank my fiancé and future wife, Leah DiNatale, who spent countless hours with me editing the content presented here and providing much needed encouragement. I would also like to acknowledge all my colleagues that I have had the privilege to work with throughout my career: Jana R. Fattic, Alanna Malone, Marsha Wallace, Gretchen Grover, Stuart Foster, Lian Zhu, Xiaoyin Zhang, Sahar T. Sadeghi, and Abdoul A. Oubeidillah. Further, I offer gratitude to those who are serving on this Committee, as they helped to refine and focus the work presented here. All of this would not be possible without funding from the U.S. Department of Defense as part of their Science, Mathematics, and Research for Transformation (SMART) Scholarship Program.

CONTENTS

| | |
|---|------|
| ABSTRACT | ii |
| DEDICATION | iv |
| LIST OF ABBREVIATIONS AND SYMBOLS | v |
| ACKNOWLEDGMENTS | vii |
| LIST OF TABLES | xi |
| LIST OF FIGURES | xiii |
| CHAPTER 1 INTRODUCTION AND HYPOTHESIS | 1 |
| Background and Motivation | 1 |
| Hypothesis and Experimental Design | 4 |
| Dissertation Research | 6 |
| Outline of Dissertation | 7 |
| CHAPTER 2 APPLICATION OF HYDRAULIC GEOMETRIES FOR LARGE SCALE WATER DEPTH AND CHANNEL ESTIMATION IN ALABAMA..... | 8 |
| Introduction | 8 |
| Review of Literature | 9 |
| Method | 15 |
| Results | 18 |

| | |
|--|----|
| Conclusions | 23 |
| Cross sectional estimation | 23 |
| Depth estimation | 31 |
| References | 38 |
| | |
| CHAPTER 3 RAPID FLOOD DAMAGE ASSESSMENT USING PUBLIC DOMAIN CADASTRAL AND ADDRESS POINT DATA WITH FUZZY LOGIC ALGORITHMS | 43 |
| Introduction | 43 |
| Review of Literature | 44 |
| Method | 49 |
| Results | 58 |
| Conclusions | 67 |
| References | 77 |
| | |
| CHAPTER 4 DEVELOPING REAL-TIME REGIONAL FLOOD DAMAGE DATASETS IN A LARGE SCALE, SHORT LEAD TIME, FLOOD DAMAGE PREDICTION ENVIRONMENT, USING FLOOD DAMAGE WIZARD | 84 |
| Introduction | 84 |
| Review of Literature | 86 |
| Method | 91 |
| Damage Estimation | 91 |
| Method of Comparison | 94 |
| Study Region | 94 |
| Results | 96 |

| | |
|---|-----|
| Scenario 1: Estimation of Market Value | 96 |
| Scenario 2: Market Value Provided | 100 |
| Conclusions | 106 |
| Future Work | 111 |
| References | 114 |
| CHAPTER 5 CONCLUSIONS | 118 |
| Research Questions and Conclusions | 118 |
| Research Question 1 | 118 |
| Research Question 2 | 119 |
| Research Question 3 | 121 |
| General Conclusions | 122 |
| Additional Research | 125 |
| REFERENCES | 129 |
| APPENDICES | 133 |
| Flood Damage Wizard Code | 133 |
| Extract_Raster_Value | 133 |
| Damage_Curve_Search | 134 |
| Data_Merger | 134 |
| Flood_Damage_Calculator | 134 |

LIST OF TABLES

| | |
|--|----|
| Table 1. Hydraulic geometries used in this research. | 17 |
| Table 2. List of Alabama streams used in this study. Each stream in this list is an unregulated stream which is necessary for the utilization of hydraulic geometries. | 18 |
| Table 3. Structural Damage Estimation Parameters. These parameters function to fill in data gaps when no local estimates are available for estimating structural value, content value, and inventory value. From Chapter 14 of the Hazus Technical Manual, excluding column 8. Occupancy Classes RES1 and RES3 have been generalized from the more expanded Hazus versions. | 57 |
| Table 4. Pertinent data provided by each county in the study. Note that Autauga County Address Points provide a condition description for part of the county..... | 61 |
| Table 5. Translation of qualitative structural condition description to yearly depreciation equivalent. These numbers are arbitrary and require additional validation. | 61 |
| Table 6. Select subset of resultant matches using Flood Damage Wizard in Tuscaloosa County, Alabama. | 62 |
| Table 7. Select probabilistic economic damage estimation from Tuscaloosa County, Alabama. | 63 |
| Table 8. Select estimated financial damages from the Tuscaloosa County, Alabama Address Point and cadastre data..... | 63 |
| Table 9. Select estimated economic damages from the Tuscaloosa County, Alabama Address Point and cadastre Data..... | 64 |
| Table 10. Flood Damage Wizard damage compared to Hazus Level I homogeneous and dasymetric analyses. | 64 |
| Table 11. Default values used by Flood Damage Wizard when the user dataset does not specify them. | 92 |
| Table 12. Data Resolution for Each Parcel Dataset Used in this Study (U.S. Census Bureau, 2015). | 96 |

| | |
|--|-----|
| Table 13. Comparison of total building damage estimates using all values except market value. Note that in cases where area with development, such as Autauga and Montgomery County, estimates in relation to Hazus are much greater..... | 98 |
| Table 14. Total damage comparison between Hazus and each of the Flood Damage Wizard (FDW) calculations. Note that this comparison omits Hazus locations where FDW was unable to find a suitable depth-damage function. | 102 |
| Table 15. Example of residential single family match in Autauga County dataset. Note that no inventory class was given to this match..... | 103 |
| Table 16. Example of residential single family match in Montgomery County dataset. Note that no inventory class was given to this match..... | 103 |
| Table 17. Example of depth-damage curve and description matches made by Flood Damage Wizard in Montgomery County. | 103 |
| Table 18. Example of questionable depth-damage curve and description matches made by Flood Damage Wizard in Autauga County..... | 104 |
| Table 19. Subset of buildings classified as RES1 in Hazus Level II database in Montgomery County, Alabama. In this instance it seems that in both Flood Damage Wizard and the Hazus data mis-classify the data in some instances. Similar circumstances occur in the Autauga County dataset..... | 110 |
| Table 20. Subset of buildings classified as COM1 in Hazus Level II database in Autauga County, Alabama. In this instance it seems that in both Flood Damage Wizard and the Hazus data mis-classify the data in some instances..... | 110 |

LIST OF FIGURES

Figure 1. Cross Sectional Profile Along the Blackwater River in Escambia County, Alabama. This cross section is part of a HEC-RAS model..... 3

Figure 2. Depth-damage function for a Meat Market. These functions typically describe the percent damage at a given depth of water for a structure. There are three general classes of depth-damage functions, structural, content, and inventory. Note that the functions for Contents (Cont) and Inventory (Inv) are identical. 4

Figure 3. Streams used to test the ability of hydraulic geometries in predicting channel dimensions. 14

Figure 4. Box Plots for HEC-RAS dimensions determining the outliers generated by HEC-RAS simulations. In one HEC-RAS cross section, an estimated width of approximately 1,000 meters is within a dataset in which the majority of estimates of width are less than 100 meters. There are several possible explanations for these outliers. The researchers remove the width and depth outliers, through box plot identification, before regression analysis presented in Figure 4. 20

Figure 5. Regression plots of width, depth, and cross-sectional area comparing estimated hydraulic geometry channel dimensions to HEC-RAS channel dimensions. The East Gulf Coastal Plain results are the top figure and the Piedmont Upland results are the bottom figure. Both physiographic sections produce approximate results with 10 times the naturalized discharge. The solid line represents a slope of 1, the dashed line is the regression line. 22

Figure 6. Relationship between HEC-RAS predicted depth and hydraulic geometry predicted depth for each multiple of gage adjusted naturalized flow (Q) for all East Gulf Coastal Plain sites (Figure 2). Note the increase in the R^2 value and the consistent slope. The solid line indicates a line with a slope of one and the dashed line is the regression line. 23

Figure 7. A cross section from the Blackwater River in Escambia County, Alabama. The solid line represents the surveyed cross-section with overbanks. The dotted line represents a rectangular hydraulic geometry estimate of the channel. 27

Figure 8. Land cover in Alabama. Note that the pervasive land cover in areas where hydraulic geometries do not match the HEC-RAS results is either developed with either agricultural or developed uses. The areas include the Highland Rim, Cumberland Plateau, and Tennessee physiographic sections..... 28

Figure 9. Flipped dimensions at ten times the naturalized flow in the Cumberland Plateau physiographic section of Alabama. Here the hydraulic geometry width is larger than those estimated by HEC-RAS and the depth is smaller than those estimated by HEC-RAS. When combined, the area of the channel approximation produced by the hydraulic geometries is approximate to that produced by HEC-RAS. The solid line represents a slope of one and the dashed line is the regression value. This result suggests that the estimates of width and depth for the Cumberland Plateau section of Alabama are opposite to those observed in other parts of the physiographic province..... 29

Figure 10. Terrain of Alabama. Note the relatively low gradient, homogeneity of the East Gulf Coastal Plain physiographic section as compared to the other physiographic sections of Alabama. This indicates that the stronger ability to predict channel geometry in this region maybe due to either the lack of relief or homogeneous relief within the relief. 30

Figure 11. Comparison of estimated depths generated by hydraulic geometries and HEC-RAS for 100-Year events in the East Gulf Coastal Plain. The results are consistent with the output generated using all multiples of gage adjusted naturalized flow from the NHDPlus Version 2 dataset. The solid line represents a slope of one and the dashed line is the regression line. 34

Figure 12. Hydraulic geometry depth functions used for each physiographic section in Alabama. Note that the functions which allow for accurate depth predictions at multiple flow regimes are those which possess a relatively high sensitivity to discharge. This characteristic may be the cause for the functions being able to accurately measure depth at relatively low and high flow regimes..... 35

Figure 13. Relationship between HEC-RAS predicted depth and hydraulic geometry predicted depth for each multiple of gage adjusted naturalized flow (Q) for all Piedmont Upland sites (Figure 2). Note that the R^2 and slope are not as significant as in the East Gulf Coastal Plain but the slope is still very consistent throughout all flow regimes. The solid line represents a slope of one and the dashed line is the regression line. 36

Figure 14. Comparison of Predicted Changes in Water Depth Based on Hydraulic Geometries and Rating Curves for USGS gage locations along the Blackwater and Conecuh Rivers. These rivers are in the East Gulf Coastal Plain. 37

Figure 15. System flowchart for flood damage assessment software (Flood Damage Wizard). 52

Figure 16. Locations of each Case Study used in this Research..... 58

Figure 17. Inundation locations used in this study (a) Pine Creek in Autauga County, Alabama (b) Black Warrior River in Tuscaloosa County, Alabama (c) Shoal Creek in Travis County, Texas. 65

Figure 18. Subset of damage locations at Address Points during 100-Year event along Pine Creek in Autauga County, Alabama. 66

Figure 19. Location where the address points falls outside the parcel polygon in Travis County, Texas. This led to a blank description for the parcel which is unusable by Flood Damage Wizard. 67

Figure 20. Comparison of Address Point data to aerial imagery. Here it is demonstrated that Address Points are, in most cases, approximate building locations. 76

Figure 21. Inverse Distance Weighted (IDW) example of how flood risk can be communicated with forecasted damage estimation. 77

Figure 23. Point total damage estimate comparison between Flood Damage Wizard maximum values and Hazus Level II analysis for each counties 500-year event. Here, the authors use no market value and there seems to be no correlation between Flood Damage Wizard and Hazus total losses. The solid line represents a slope of one and the dashed line is the regression line. When the structural age is unknown, as in Autauga and Montgomery Counties, the maximum value is the full replacement value and Flood Damage Wizard overestimates loss. When structural age is known, as in Montgomery County, Flood Damage Wizard underestimates value and resulting flood loss. 99

Figure 24. Point total damage estimate comparison between Flood Damage Wizard maximum estimate and Hazus Level II analysis for each counties 500-year event. Here, the use of market value occurs and two of three instances seem to provide similar results. The solid line represents a slope of one and the dashed line is the regression line. 105

Figure 25. Google Street View® of Point 131 in Montgomery County. The parcel data classifies this as a service shop, the Hazus Level II building inventory classifies the site as single family residential, and from Google Street View®, the site appears to be a church..... 111

CHAPTER 1 INTRODUCTION AND HYPOTHESIS

In this chapter, the researcher introduces the needs for research in this area. Following a summary of the background information, this chapter presents the research design and a list of research questions to be answered by the dissertation. This chapter also presents the structure of the dissertation manuscript.

Background and Motivation

Hydraulic and damage assessment of riverine floods traditionally occurs in static, long-term planning environments. Hydraulic and damage assessment are critical in understanding the impacts of regional flooding events. Recent advances in continental and global scale hydrology (Oubeidillah et al., 2014; Maidment, 2014; Archfield et al., 2015; Sood and Smakhtin, 2015) make possible the ability to predict discharge at all locations in the coterminous United States (CONUS). These advances are becoming operational as the National Flood Interoperability Experiment (NFIE) creates the foundation for the National Water Model at the National Water Center (NWC). Theoretically, an extension from discharge outputs the National Water Model will produce into hydraulic and damage assessment of floods is possible. However, a number of questions concerning large scale, short lead-time hydraulic and flood damage analyses require further investigation.

First, large scale, CONUS-wide hydraulic analyses would require cross-sectional profiles of the river channel and floodplain and/or a profile of the entire channel and floodplain. Cross sections are necessary for one-dimensional hydraulic modeling. Figure 1 is an example of a

cross section used in the one-dimensional hydraulic model, Hydrologic Engineering Center's River Analysis System (HEC-RAS). Figure 1 illustrates that a typical cross section contains both the floodplain and the stream channel. Cross sections such as Figure 1 are available only in limited locations across the CONUS (Hodges, 2013). Fully developed bathymetric characteristics are typically only found for large, economically viable rivers, such as those maintained by the U.S. Army Corps of Engineers (USACE). Full profiles of bathymetric characteristics are necessary in two- and three-dimensional modeling (Merwade et al., 2008). Thus, any attempt at large scale hydrodynamic modeling would require a means by which to fill data gaps in locations where cross sections or channel information is unavailable. Hydrodynamic models or other means of estimating depth-discharge relationships are necessary to estimate flood impact spatially and fiscally, as depth-damage relationships are the preeminent means of flood damage estimation (Eckstein, 1965; Grigg and Helweg, 1975; Das and Lee, 1988; Yang and Tsai, 2000; Freni et al., 2010). Figure 2 is an example of a depth-damage relationship for a meat market. Here, the figure demonstrates that there are three general classes of depth-damage function, structural, content, and inventory. In general, the functions attribute a percent of damage to each component of a building given a depth of water.

Once an accurate representation of flood inundation is available, it is possible to estimate the resulting damage through regional building inventories. Investigators assemble geospatial building inventories that describe, point by point, the use and classification of buildings for a local region and the depth-damage function for these buildings (Ding et al., 2008; Tate et al., 2015). These datasets are typically referred to as building inventories. In the contemporary, these datasets take large amounts of capital and manpower in order to assemble them effectively (Ding et al., 2008; FEMA, 2012; Tate et al., 2015). These datasets are available for only select

locations. The use of building inventories in the United States usually revolves around the assemblage of Flood Risk Reports (FEMA 2014). Thus, a means to quickly develop a building inventory is necessary to rapidly assess flood damage in real-time. Prediction of floods and the resulting damage can provide a range of benefits: 1) a fundamental first step in understanding economic impact of a flood event; 2) a mechanism for flood risk mapping; 3) an ability to proactively decrease economic impact to floods and; 4) a means to more effectively allocate emergency funds prior to flooding events (Oliveri and Santoro 2000; Merz et al. 2010; Pappenberger et al. 2015).

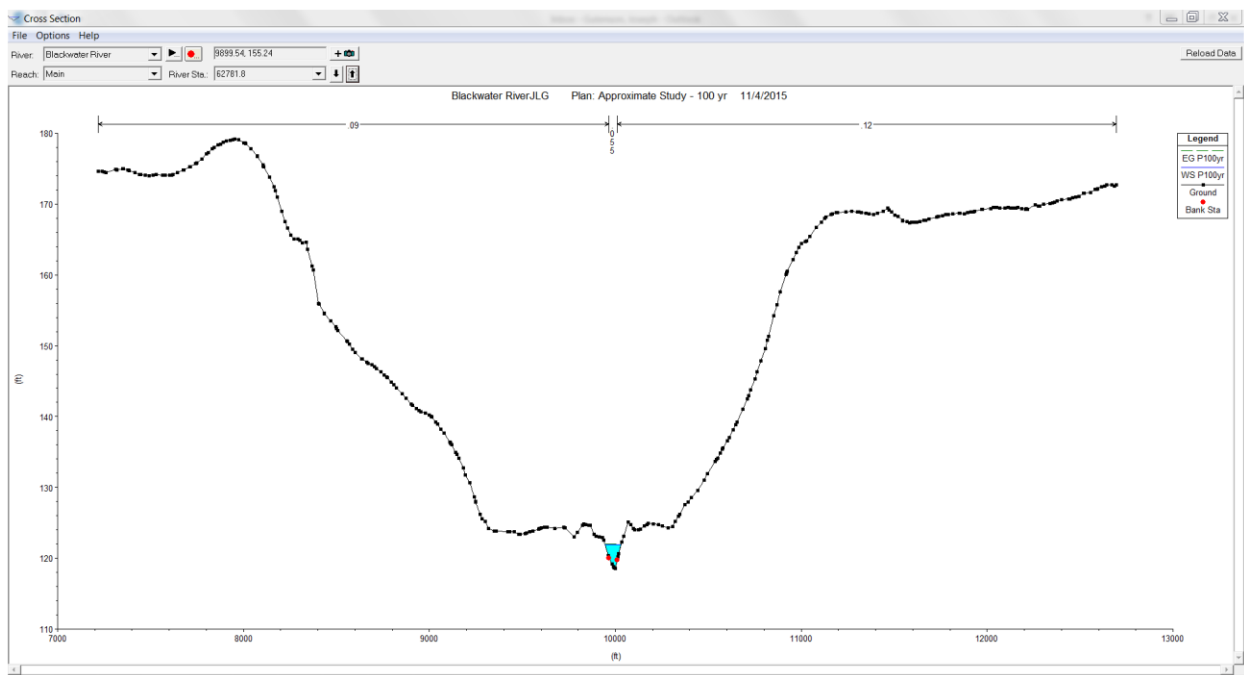


Figure 1. Cross Sectional Profile Along the Blackwater River in Escambia County, Alabama. This cross section is part of a HEC-RAS model.

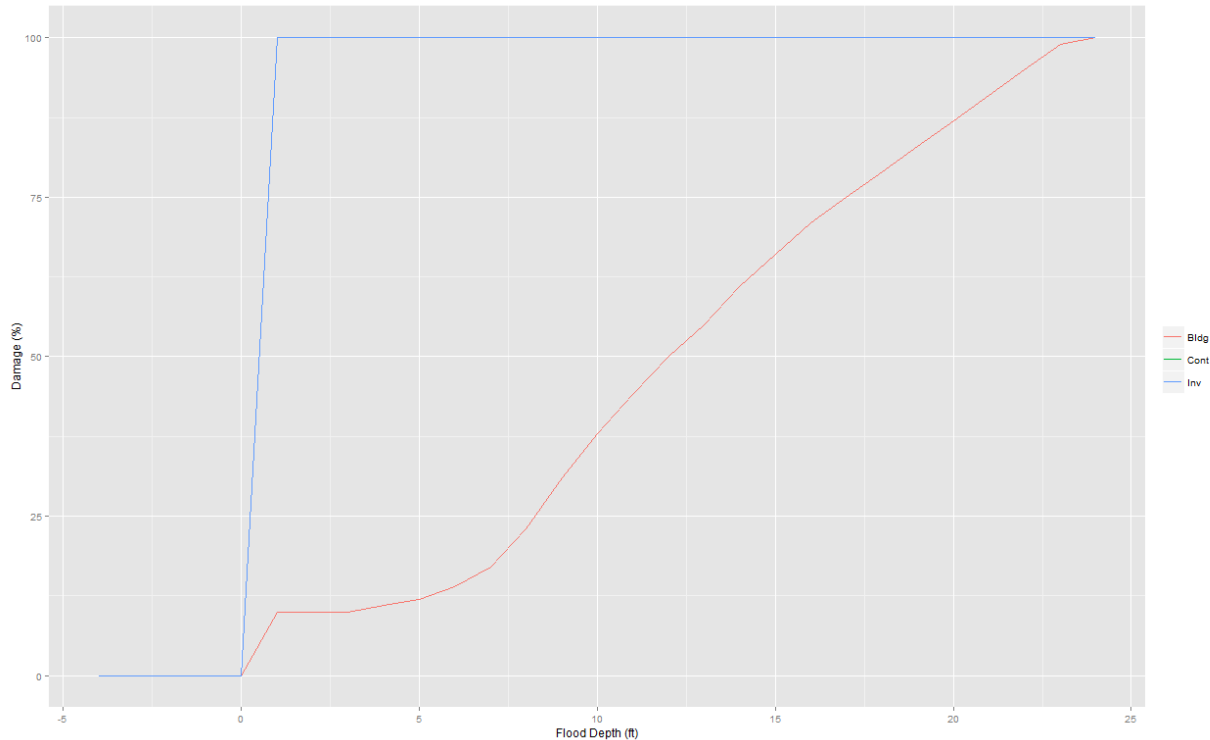


Figure 2. Depth-damage function for a Meat Market. These functions typically describe the percent damage at a given depth of water for a structure. There are three general classes of depth-damage functions, structural, content, and inventory. Note that the functions for Contents (Cont) and Inventory (Inv) are identical.

Hypothesis and Experimental Design

The first hypothesis of this research is that hydraulic geometries are capable of accurately estimating channel dimensions for both developing synthetic cross-sections and composite channel geometry. Leopold and Maddock (1953) first introduced hydraulic geometries with their fundamental work for the United States Geological Survey (USGS). Hydraulic geometries describe a natural relationship between the bankfull discharge of a stream and its width, depth, and cross-sectional area. The second hypothesis is that the hydraulic geometry methodology can be employed to estimate depth, specifically during conditions above and below bankfull discharge. The research tests these potential applications by comparing HEC-RAS predicted

flow characteristics (depth, width, and cross-sectional area) against those derived from hydraulic geometries. The researcher identifies the location of HEC-RAS cross sections along the NHDPlus Version 2 streamlines in five physiographic regions in Alabama. The HEC-RAS simulations use gage adjusted naturalized flow from the NHDPlus Version 2 dataset as streamflow input. Gage adjusted naturalized flow is considered mean annual streamflow (Bondelid, 2014).

Emergency managers across CONUS maintain Address Points, which are vector spatial data, approximate to a building's location. Most U.S. County tax assessment offices produce georeferenced cadastral data. To varying degrees, these parcel data describe building characteristics of structures within the parcel. The third hypothesis is that using a combination of Address Point and cadastral information can provide a feasible estimate of regional flood damage. Joining Address Point data with cadastral data offers the ability to rapidly develop building inventories. The researcher asserts that a fuzzy logic text matching schema can effectively determine a buildings depth-damage function by comparing the text description of the function with each cadastral use description. The author evaluates a methodology that relies upon Address Point and parcel datasets in three U.S. counties to forecast damage estimates to buildings in CONUS.

Further, the researcher investigates what information is necessary to approximate detailed flood studies such as those produced in Federal Emergency Management Agency (FEMA) Flood Risk Reports by Hazus, a software by FEMA. The researcher hypothesizes that the fuzzy logic algorithm can approximate a detailed study with the use of data available within a cadastral dataset. This research analyzes which cadastral inputs help planners and researcher to best understand flood damage by comparing the damage results generated by a detailed Hazus

analysis in the Upper Alabama Watershed Flood Risk Report (FRR) near Montgomery, Alabama to those approximated using cadastral inputs. The research uses the 10-, 25-, 50-, 100-, and 500-year return period events along the Upper Alabama River in three counties within the Upper Alabama watershed. These counties are Autauga, Lowndes, and Montgomery. This generates 15 samples for comparison. The researcher considers both composite and individual total damage in these scenarios.

Dissertation Research

The primary objective of this research is to characterize the viability and the input necessary to perform real-time flood damage assessment in a short lead-time environment. Thus, a means to estimate depth-discharge relationships, construct building inventories, and estimate flood damage is under investigation. Specifically the research examines the following questions:

- 1) Can the use of hydraulic geometry determine channel dimensions of streams where no cross-sectional or bathymetric data is available? Are hydraulic geometries useful in estimating synthetic depth-discharge relationships on streams which are above and below bankfull discharge?
- 2) What method of linking the cadastral description of a building to the appropriate depth-damage function and developing a building inventory in real-time is possible?
- 3) What inputs into a building inventory are necessary to approximate a detailed flood damage analysis?

Outline of Dissertation

This dissertation consists of three articles, each building toward the support of a singular goal, a method of rapidly estimating flood water depth and translating water depth into resulting building damage. These articles compose Chapters 2, 3, and 4, respectively.

Chapter 2 introduces how the hydraulic geometry concept can estimate channel geometry for hydraulic modeling through approximating bankfull discharge. The research also considers how hydraulic geometries may approximate depth of flow for discharge above and below bankfull. The research tests these assertions in each of the five physiographic sections of Alabama.

Chapter 3 discusses the implementation of a fuzzy logic framework for predicting flood damage. This research highlights the functionality of the system and the inputs that it utilizes. The research applies the methodology in Autauga County, Alabama, Tuscaloosa County, Alabama, and Travis County, Texas.

Chapter 4 analyzes what inputs are important for the fuzzy logic framework to be able to approximate detailed flood damage studies. The study uses detailed Hazus Level II datasets from the Upper Alabama FRR. These Level II datasets are from three counties in the Upper Alabama Watershed, specifically Autauga County, Lowndes County, and Montgomery County.

Chapter 5 highlights the results of these experiments, summarizes the findings, and provides recommendation for future work.

CHAPTER 2 APPLICATION OF HYDRAULIC GEOMETRIES FOR LARGE SCALE WATER DEPTH AND CHANNEL ESTIMATION IN ALABAMA

Introduction

This study investigates the use of hydraulic geometries within the existing framework of the National Flood Interoperability Experiment (NFIE) for the state of Alabama. Hydraulic geometry coefficients in hydrological applications include three avenues of usage: 1) an alternative for estimating channel geometries for cross-sections or hydraulic modeling; 2) the development of discharge-depth rating curves where none exist; 3) an estimate of flood inundation depths in data-poor river reaches. We test these potential applications by comparing HEC-RAS predicted flow characteristics (depth, width, and cross-sectional area) against those derived from hydraulic geometries. We identify the location of HEC-RAS cross sections along the NHDPlus Version 2 streamlines in five physiographic regions in Alabama and run HEC-RAS simulations using gage adjusted naturalized flow within the NHDPlus Version 2 dataset as streamflow input. Results indicate that all three potential use scenarios are possible using hydraulic geometry on streams within the East Gulf Coastal Plain and Piedmont Upland physiographic sections of Alabama, where anthropogenic influence on the stream channel is minimal. Local geologic conditions may play a role in the effectiveness of the hydraulic geometry coefficients. Evidence supports that an approximate water depth estimate is possible in these physiographic sections and possibly, other regions with similar geomorphic characteristics.

Review of Literature

Stream channel water transport capacity is a key parameter for flood inundation and instream hydrodynamic processes modeling and analysis. Recent advances in large-scale hydrological modeling now provide highly spatially and temporally resolved streamflow predictions at continental (Oubeidillah et al., 2014; Archfield et al., 2015) and global scales (Sood and Smakhtin, 2015). Most notably, the modeling framework developed by Maidment (2014) provides volumetric runoff calculations at near real-time temporal resolution and are routed as streamflow for nearly 2.7 million river reaches over the coterminous United States (CONUS). The Maidment (2014) modeling framework was part of the National Flood Interoperability Experiment (NFIE) which advanced a new flood prediction system for the United States. In order to translate the model streamflow predictions to flooding potentials, the relationship between streamflow and water depth needs to be resolved. One approach undertaken to quantify the relationship between streamflow and water depth is through channel cross-section analysis based on digital elevation models (DEMs). National coverage of DEMs at 10 x 10 m resolution are available through the U.S. Geological Survey (USGS 2015) and acquisition of higher resolution DEMs is possible using aerial LiDAR and sonar bathymetric datasets. However, complete bathymetric datasets are typically only available along large, economically viable rivers, such as those maintained by the U.S. Army Corps of Engineers (USACE).

An alternative approach for estimating streamflow-depth relationship is the use of hydraulic geometry relationships. Leopold and Maddock (1953) first introduced this method with their fundamental work for the USGS. Since that time, numerous applications based on these concepts sparked further research. Merwade, for example, published a number of tutorials for estimating bathymetric datasets and expedite hydrodynamic modeling based on hydraulic geometry

(Merwade 2014). In particular, his tutorial for the River Channel Morphology Model (RCMM) discusses the use of hydraulic geometry relationships to estimate channel depth and width at user-specified cross-sections where no ground measurements have been made (Merwade, 2006). Hydraulic geometry relationships have become a fixture in stream restoration activities (Smoot et al., 2014) and are commonly used as input parameters in hydrological models, such as the Soil and Water Assessment tool (SWAT; Arnold et al., 1998) and Hydrologic Simulation Program-Fortran (HSPF; Bicknell et al., 1997).

Though hydraulic geometry application extends to velocity, Manning's roughness factor, and other values, for the purposes of this research, they take the form of the following power functions:

$$w = aQ^b ; d = cQ^f ; a = kQ^m \quad (1)$$

where w is mean width, d is mean depth, a is the mean cross sectional area, and Q is the bankfull discharge of a river reach. Parameters a , c , k , b , f , and m are unitless coefficients empirically derived through observation. The bankfull discharge, Q , in these relationships represents an unregulated flow. Thus, the intended usage for hydraulic geometries is for streams under natural flow regimes (Leopold 1953). The hydraulic geometry parameters demonstrate a significant relationship with the physiographic region of the stream (Bieger et al. 2015). The USGS, U.S. Fish and Wildlife Service (FWS) (McCandless 2003a; 2003b), U.S. National Resource Conservation Service (NRCS) and others publish numerous studies throughout the U.S. quantifying the coefficients listed above. Stewardson (2005) provides an example of usage in other countries. Alternatively, many compute a relationship between a , c , k , b , f , m , and drainage area. These drainage area relationships have proven to be less reliable than discharge (Bieger et al., 2015).

Mohamoud and Parmar (2006) utilized USGS regional regression equations to estimate mean annual streamflow and they used this streamflow to calculate hydraulic geometries for wadeable streams in the Mid-Atlantic region of the United States. That experiment demonstrates strong predictive capabilities when compared to the U.S. Environmental Protection Agency (USEPA) Environmental Monitoring and Assessment Program (EMAP) recorded depth. Hence, Mohamoud and Parmar (2006) provide a precedence of using an estimated mean annual streamflow in hydraulic geometry relationships. Davies et al. (2007) applied a methodology rooted in the hydraulic geometry relationships in estimating channel geometry. Through examination of orthoimagery, they empirically developed regression equations using log-transformed variables. Validation of this method occurred through comparison with field observations. The results demonstrate high correlation with field observations. An additional study by Magirl and Olsen (2009) examined the practicality of estimating hydraulic geometries to determine stream navigability in the state of Washington. That study utilized the mean annual flow from the NHDPlus dataset. The results of the Magirl and Olsen (2009) study were mixed. However, they concluded that their methodology is a good predictor of stream navigability. Hodges (2013) also asserts that hydraulic geometries may find use in large scale synthetic channel approximation.

As part of the NFIE, the NFIE-Geo framework provides a comprehensive geospatial dataset for hydrologic modeling. This composite dataset rectifies the piecemeal nature of data assimilation for hydrologic inquiry and provides the ability to easily download the NHDPlus Version 2 dataset. The NHDPlus Version 2 dataset of 2.67 million stream segments is within NFIE-Geo as 12 NFIE-Hydro Regions (Maidment, 2015). The NHDPlus datasets contain an estimate of gage adjusted, naturalized flow using the Enhanced Runoff Method (EROM). The

EROM method makes use of the 21 Vector Processing Units (VPUs) which comprise the NHDPlus Version 2 dataset. This flow is an approximation to mean annual streamflow for each reach NHDPlus Version 2 reach. Gage adjusted naturalized flow represents the best estimate within the NHDPlus Version 2 dataset. The basis of improvement in the naturalized flow estimates is the number of gages available to adjust the estimated flow (Bondelid 2014).

Physiographic regions provide a taxonomy of the landforms of Earth into areas of homogeneous geomorphology (Fenneman 1917). Alabama falls within three physiographic divisions: the Appalachian Highlands, Atlantic Plain, and Interior Plains. These three divisions subdivide into five sections: the East Gulf Coastal Plain, Cumberland Plateau, Tennessee, Piedmont Upland, and Highland Rim (Figure 3). Limestone bedrock and moderate relief characterize the Highland Rim section. Sandstone and Shale bedrock and moderate relief characterize the Cumberland Plateau. Valley and Ridge formations dominate the topography of the Tennessee section with valleys comprised of carbonate and shale and ridges which are predominately sandstone or chert. Moderate relief and crystalline bedrock characterize the Piedmont Upland physiographic section. Deep sediment deposits and low to moderate relief form the East Gulf Coastal Plain province which envelopes most of Alabama (Raymond et al. 1988).

In this study we investigate the potential applicability of hydraulic geometries in large scale hydrologic and hydraulic modeling applications. This is the first comprehensive study of hydraulic geometries for the entire state of Alabama. We first investigate whether hydraulic geometries can be used to approximate channel characteristics for ungagged, non-surveyed streams in Alabama. These channel characteristics may be functional in one-dimensional hydraulic modeling where no detailed cross sectional channel information is currently available

when combined with existing DEM data. Further, the results may generate a continuous, synthetic channel network, in data poor environments, which are necessary for two- and three-dimensional hydraulic models (Merwade et al. 2008). The results may also provide a means of altering DEMs for hydrologic routing, such as the Variable Infiltration Capacity (VIC) model instance developed by Oubeidillah et al. (2014). Next, we investigate the extended role of the hydraulic geometries in estimating water depth in both bankfull and overbank (flood) discharge scenarios. As the intended usage of hydraulic geometries is for only bankfull discharge, nothing found in the literature investigates other flow scenarios. The effectiveness of accurate flow-depth relationship approximation with this approach is novel and useful for generating depth-discharge rating curves for one-dimensional hydraulic models and may function as an approximator of inundation extent.

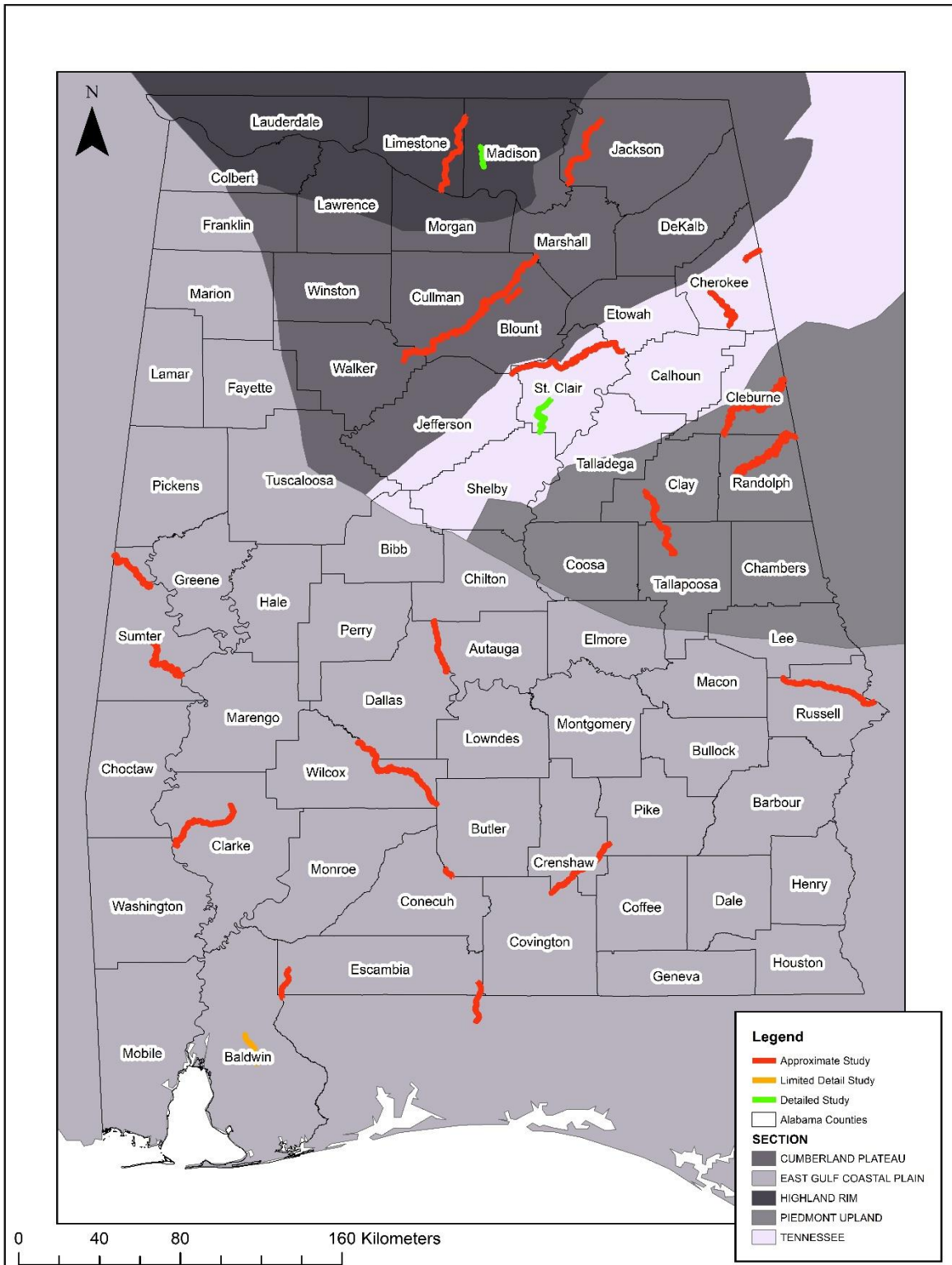


Figure 3. Streams used to test the ability of hydraulic geometries in predicting channel dimensions.

Method

We investigate the usability and robustness of hydraulic geometry relationships for estimating water levels and channel geometry in Alabama. The Alabama study area includes over 200,000 stream segments in the NHDPlus Version 2 dataset, or approximately 7.5% of the NHDPlus Version 2 dataset. The NHDPlus Version 2 gage adjusted naturalized streamflow provides a basis for estimating bankfull discharge. We compare the depth, width, and area estimates generated by hydraulic geometries to surveyed cross sections in 25 Hydrologic Engineering Center River Analysis System (HEC-RAS) models (Table 2) produced by the Alabama Department of Economic and Community Affairs (ADECA). The use of HEC-RAS allows us to compare our calculations under multiple streamflow regimes. The United States Geological Survey (USGS) considers all streams used in this study as unregulated (USGS, 2013).

The HEC-RAS dataset represents each of the five physiographic sections of Alabama (Figure 3). Each model is either an approximate study or detailed study. This cataloging index describes whether the HEC-RAS models include inline structures (i.e., culverts) and/or stream crossings. Detailed studies also have more stringent constraints on the topographic data used to build the hydraulic model (Phillip Hicks, Alabama Department of Economic and Community Affairs, personal correspondence).

No published hydraulic geometry expressions relating discharge to channel width, depth, and cross sectional area exist for the state of Alabama. We use approximate relationships from McCandless (2003a; 2003b), Mohamoud and Parmar (2006), the NRCS (NRCS), and Brockman (2010) instead. These constitute approximate coefficients and originate in other regions of the United States. The coefficients utilized correlate with the same physiographic province of each

physiographic section in Alabama. Research conducted by Johnson and Fecko (2008) justifies the use of these coefficients from other areas, but in the same physiographic province. Table 1 lists these coefficients. Brockman (2010) does not provide an area estimate; we use the approximate depth and width to estimate a rectangular cross sectional area for streams within the Highland Rim physiographic section. Each equation expresses discharge (Q) in cubic feet per second, except the estimates of Mohamoud and Parmar (2006), which lists Q in cubic meters per second.

We estimate Q in each channel using gage-adjusted naturalized flow for each river reach. Because naturalized flow is equivalent to mean annual discharge, we use multiples of 2, 4, 6, 8, 10, 100, and 1,000 to investigate high flow regimes and to see if they yield accurate hydraulic geometries. We assume that 100 and 1,000 multiples of naturalized flow are equivalent to flood stage within each stream. This assumption stems from comparison of 100-Year flows within each of the HEC-RAS models to the gage adjusted naturalized flow multiplied by 100 and 1000. For example, the average 100-Year Event in the dataset is approximately 522 cubic meters per second. The average of the naturalized flows multiplied by 100 is approximately 1,962 cubic meters per second. The determination of these HEC-RAS flows comes from either USGS regional regression equations in approximate models or, in detailed studies, flows produced by softwares such as the U.S. Army Corps of Engineers Hydrologic Engineering Center Hydrologic Modeling System (HEC-HMS).

We compare hydraulic geometry results to the HEC-RAS results by overlaying the cross sections with the NHDPlus Version 2 dataset. The multiples of Q from each of the NHDPlus flow lines pair with their spatially coincidental cross section. These flows are inputs into steady-state HEC-RAS simulations. We then extract hydraulic depth, channel width, and channel cross-

sectional area from the HEC-RAS results and compare them to the width, depth, and cross-sectional area of the hydraulic geometry calculations. The comparison mechanism is general least squares linear regression of each HEC-RAS output on each corresponding hydraulic geometry output.

This research assumes that the HEC-RAS model accurately calculates depth, width, and cross-sectional area for a given flow. Horritt and Bates (2002) find that HEC-RAS performs well in capturing flood extents when provided a high resolution DEM. In fact, in the Horrit and Bates (2002) research, they find that HEC-RAS performs as well as more sophisticated two-dimensional models under certain circumstances.

Table 1. Hydraulic geometries used in this research.

| Physiographic Section | Width Equation | Depth Equation | Cross-Sectional Area Equation | Source |
|--------------------------------|-----------------------|-----------------------|--------------------------------------|----------------------------|
| East Gulf Coastal Plain | $2.82Q^{0.47}$ | $0.32Q^{0.4}$ | $0.89Q^{0.87}$ | McCandless (2003a) |
| Piedmont Upland | $1.6Q^{0.49}$ | $0.28Q^{0.391}$ | $0.458Q^{0.878}$ | NRCS |
| Tennessee | $2.65Q^{0.47}$ | $0.3Q^{0.33}$ | $0.79Q^{0.8}$ | McCandless (2003b) |
| Cumberland Plateau | $10.21Q^{0.48}$ | $0.29Q^{0.24}$ | $3.26Q^{0.67}$ | Mohamoud and Parmar (2006) |
| Highland Rim | $3.522Q^{0.4406}$ | $0.2525Q^{0.3898}$ | N/A | Brockman (2010) |

Table 2. List of Alabama streams used in this study. Each stream in this list is an unregulated stream which is necessary for the utilization of hydraulic geometries.

| Stream | County | Physiographic Section | Model Type |
|--------------------------------|----------------------------|------------------------------|-------------------|
| Blue Springs Creek | Blount County, Alabama | Cumberland Plateau | Approximate |
| Mulberry Fork | Cullman County, Alabama | Cumberland Plateau | Approximate |
| Paint Rock River | Jackson County, Alabama | Cumberland Plateau | Approximate |
| Blackwater River | Escambia County, Alabama | East Gulf Coastal Plain | Approximate |
| Conecuh River | Crenshaw County, Alabama | East Gulf Coastal Plain | Approximate |
| Mulberry Creek | Autauga County, Alabama | East Gulf Coastal Plain | Approximate |
| Noxubee River | Sumter County, Alabama | East Gulf Coastal Plain | Approximate |
| Perdido River | Escambia County, Alabama | East Gulf Coastal Plain | Approximate |
| Pine Barren Creek | Wilcox County, Alabama | East Gulf Coastal Plain | Approximate |
| Satilpa Creek | Clarke County, Alabama | East Gulf Coastal Plain | Approximate |
| Sepulga River | Butler County, Alabama | East Gulf Coastal Plain | Approximate |
| Styx River | Baldwin County, Alabama | East Gulf Coastal Plain | Limited Detailed |
| Uchee Creek | Russell County, Alabama | East Gulf Coastal Plain | Approximate |
| Indian Creek | Madison County, Alabama | Highland Rim | Detailed |
| Limestone Creek | Limestone County, Alabama | Highland Rim | Approximate |
| Hillabee Creek | Tallapoosa County, Alabama | Piedmont Upland | Approximate |
| Little Tallapoosa River | Randolph County, Alabama | Piedmont Upland | Approximate |
| Tallapoosa River | Cleburne County, Alabama | Piedmont Upland | Approximate |
| Big Canoe Creek | St. Clair County, Alabama | Tennessee | Approximate |
| Chatooga River | Cherokee County, Alabama | Tennessee | Approximate |
| Kelly Creek | St. Clair County, Alabama | Tennessee | Detailed |
| Terrapin Creek | Cherokee County, Alabama | Tennessee | Approximate |

Results

The regression analysis yielded robust and consistent relationships for two out of the five physiographic regions: the East Gulf Coastal Plain and Piedmont Upland (Figure 2). Hydraulic geometries, after removing all cross sections where HEC-RAS channel dimensions are outliers from consideration, provides good approximation of flow characteristics in these physiographic sections (Figure 5). In both sections, ten times the naturalized flow seems to provide an approximate bankfull discharge. Figure 4 demonstrates an example of the magnitude of these width, depth, and cross-sectional area outliers in the data. The Discussion section of this research highlights possible explanations for the outliers in the data.

Under each flow scenario examined, the results indicate a positive correlation between HEC-RAS depths and hydraulic geometry depths. Further, R^2 increases in higher flow regimes. Figure 6 demonstrates the positive correlation, increasing R^2 and Slope coefficients of these linear regressions range from 0.77-1.1, indicating a near 1:1 relationship between the HEC-RAS hydraulic depth and the hydraulic geometry mean depth in the East Gulf Coastal Plain. Figure 6 and Figure 13 demonstrate that it is possible under normal and high flow scenarios to estimate the resulting water depth reasonably well using only the hydraulic geometry parameters. The consistency of the slope term suggests that the relationship between HEC-RAS depths and hydraulic geometries does not change between low and high flow events. This relationship also occurs within the Piedmont Upland sample, though the R^2 and slope indicate a weaker relationship than that which is found in the East Gulf Coastal Plain, demonstrated in Figure 13.

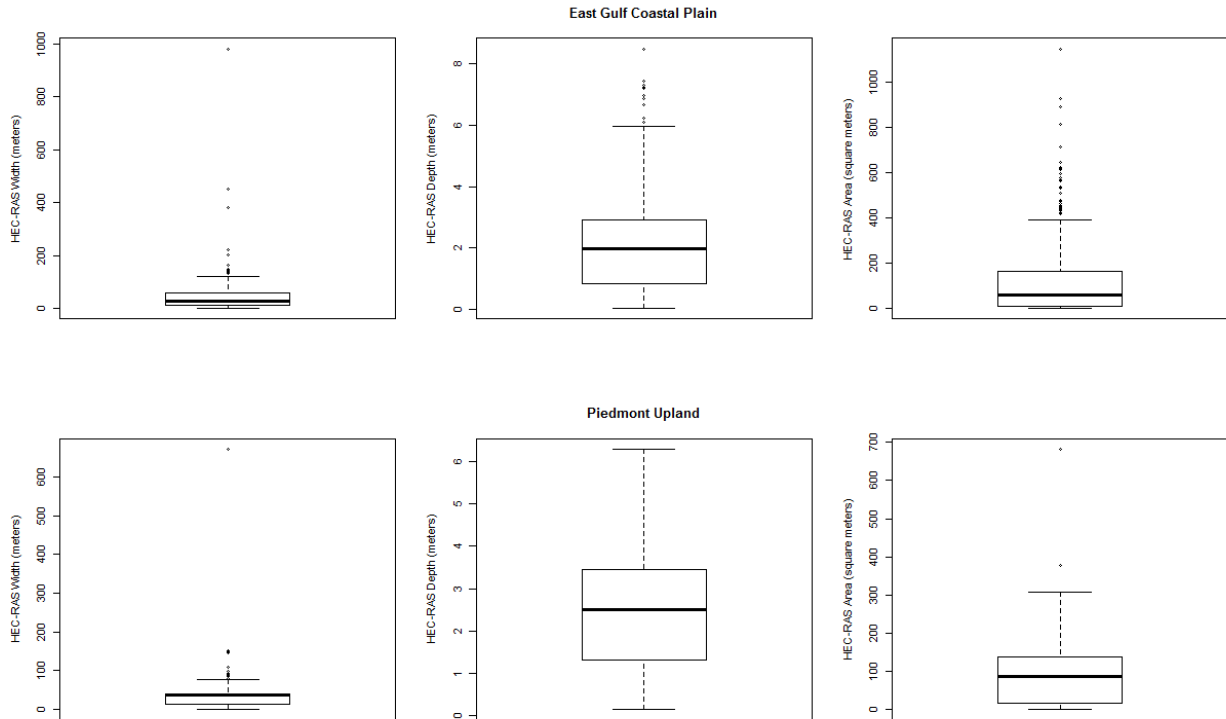
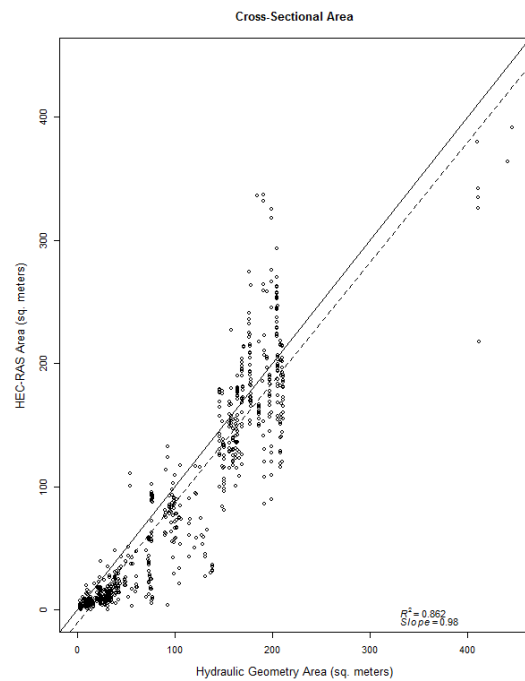
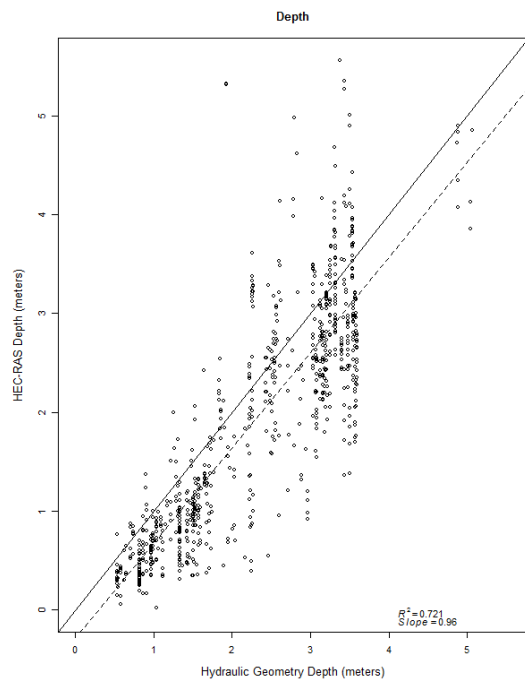
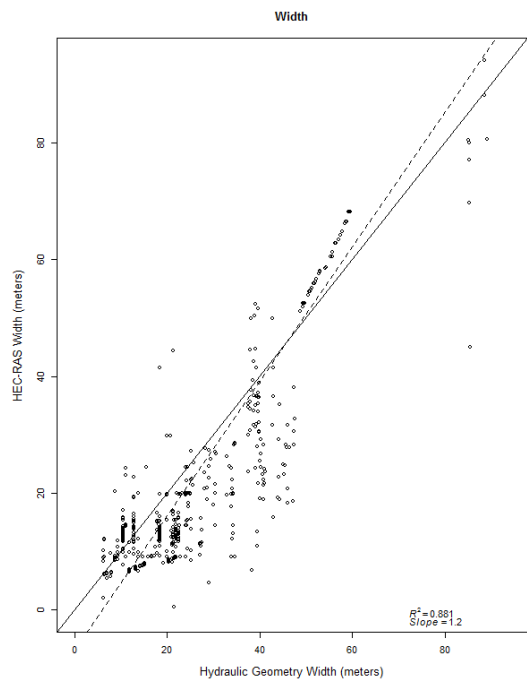


Figure 4. Box Plots for HEC-RAS dimensions determining the outliers generated by HEC-RAS simulations. In one HEC-RAS cross section, an estimated width of approximately 1,000 meters is within a dataset in which the majority of estimates of width are less than 100 meters. There are several possible explanations for these outliers. The researchers remove the width and depth outliers, through box plot identification, before regression analysis presented in Figure 4.



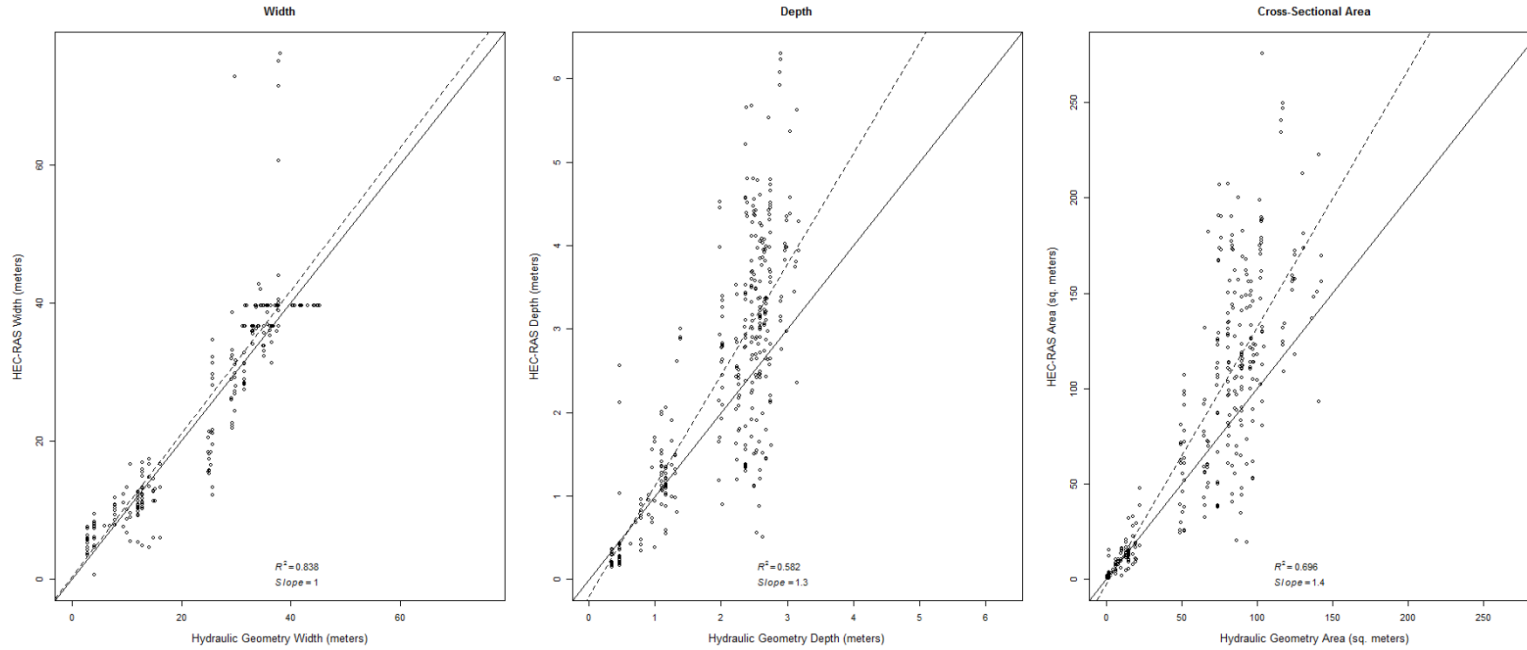


Figure 5. Regression plots of width, depth, and cross-sectional area comparing estimated hydraulic geometry channel dimensions to HEC-RAS channel dimensions. The East Gulf Coastal Plain results are the top figure and the Piedmont Upland results are the bottom figure. Both physiographic sections produce approximate results with 10 times the naturalized discharge. The solid line represents a slope of 1, the dashed line is the regression line.

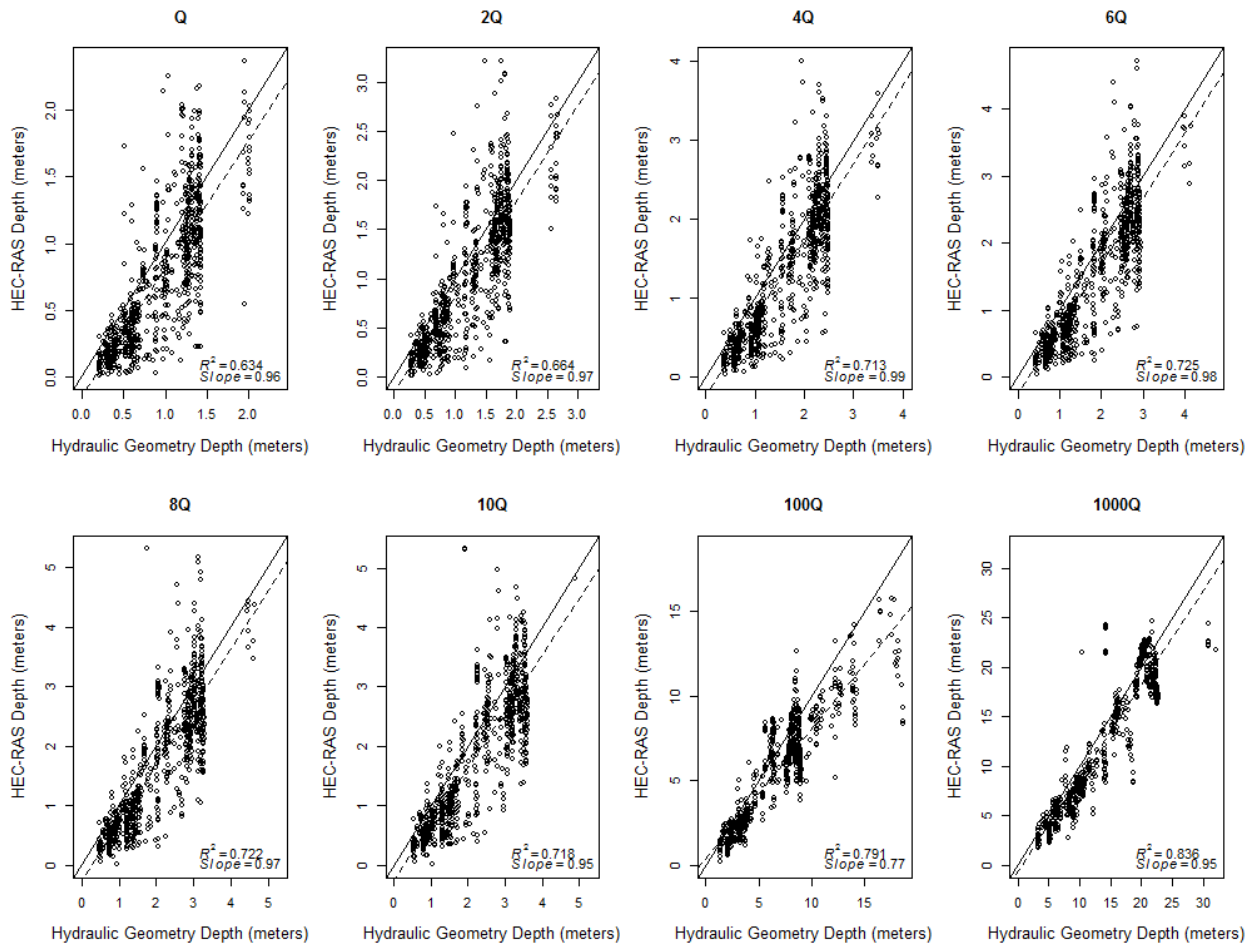


Figure 6. Relationship between HEC-RAS predicted depth and hydraulic geometry predicted depth for each multiple of gage adjusted naturalized flow (Q) for all East Gulf Coastal Plain sites (Figure 2). Note the increase in the R^2 value and the consistent slope. The solid line indicates a line with a slope of one and the dashed line is the regression line.

Conclusions

Cross sectional estimation

The challenge of estimating reasonably accurate channel morphology data is partially solved using hydraulic geometries in this study. With the removal of significant HEC-RAS channel dimension outliers, channel depth, width, and cross-sectional area estimates were consistently and simultaneously approximate to HEC-RAS cross sections in the East Gulf Coastal Plain and

Piedmont Upland physiographic sections. For example, Figure 7 demonstrates that a rectangular channel estimated using hydraulic geometries is roughly equivalent to a surveyed HEC-RAS cross-section along the Blackwater River in Escambia County, Alabama. However, outliers in the HEC-RAS outputs seem to demonstrate the tremendous influence of anthropogenic alteration on stream channels. The channels themselves may be directly altered through dredging or widening by humans or through inline structures, such as culverts. For example, historically the Paint Rock River in the Cumberland Plateau physiographic section has undergone historic channelization and dredging (Freeman and Northcutt, 2014). In this stream, HEC-RAS channel dimensions, particularly depth, are greater and uncorrelated with the hydraulic geometry dimensions. These unexplained dimensional characteristics of the stream likely reflect the anthropogenic influences on this stream. Further, as this stream is one of only three streams from the Cumberland Plateau in this study, it skews the resulting correlations.

The work of Hardison et al. (2009), McBride and Booth (2005) and Walter and Merritts (2008) discusses the impact watershed scale development can have on the physical characteristics of stream channels. Figure 8 displays land cover for the state of Alabama. Figure 8 seems to indicate that areas where the hydraulic geometries demonstrate little to no correlation with HEC-RAS outputs are pervasively developed, particularly in the Highland Rim physiographic section. This link might explain why the hydraulic geometries are inconsistent with the HEC-RAS cross sections in these regions.

Another possibility is that some of the flow regimes in this study do not reflect bankfull discharge in the Highland Rim, Cumberland Plateau, and Tennessee physiographic sections. According to Bent and Waite (2013), Castro and Jackson (2001), and Rosgen (1994) bankfull discharge is approximately equivalent to a 1.5 year recurrence interval. The assumption that a

particular multiple of gage adjusted naturalized flow approximates bankfull discharge is problematic.

This study utilizes hydraulic geometries coefficients in Alabama based on studies from different parts of the United States, the coefficients may not accurately reflect conditions in Alabama. For example, Figure 9 demonstrates that though hydraulic geometries overestimate width and underestimate depth in the Cumberland Plateau, the channel area estimates are approximate to HEC-RAS outputs at ten times the naturalized flow. This suggests that the hydraulic geometry relationship in this physiographic province are different than the relationships found in other parts of the province.

The discrepancy between HEC-RAS and hydraulic geometry estimates in the Highland Rim, Cumberland Plateau, and Tennessee physiographic sections may also be explained in the topography of these regions. The Highland Rim and Piedmont Upland and the Tennessee and Cumberland Plateau physiographic regions has relatively moderate topography (Figure 10). The East Gulf Coastal Plain physiographic section displays less rugged terrain than all of the four other physiographic sections. Streams in the Southern Coastal Plain typically have broad alluvial features and low gradients (Hupp, 2000). Further, the Cumberland Plateau and Tennessee physiographic sections also possess intermittent expanses of low gradient topography. These changes in relief likely lead to variable geometric properties of streams in the Cumberland Plateau and Tennessee physiographic sections. Thus, if the hydraulic geometry coefficients do capture this variability, the hydraulic geometries cannot perform well. Though Hupp (2000) asserts that streams in Coastal Plain regions do not typically conform to hydraulic geometries, this research indicates that hydraulic geometry coefficients provide reasonable and consistent depth estimates in such areas.

Though all physiographic sections in this study have at least 164 cross sections under comparison, these cross sections represent only 2-4 streams. Thus, autocorrelation artifacts from one stream will weigh heavily on composite results in that physiographic section. In all regions where the hydraulic geometries do not correlate with the HEC-RAS outputs, this study uses only 2-4 streams. Thus, the number of streams used in each physiographic section may have skew the results.

An alternative means of generating channel estimations is through remote sensing applications. Mersel et al. (2013) discuss the use of the Surface Water and Ocean Topography (SWOT) satellite (planned to be launched in 2020) in estimating water topography. Yoon et al. (2012) also explore a method of using SWOT data in estimating river bathymetry. These studies suggest that the use of SWOT in estimating unknown bathymetry is promising.

The use of LiDAR-based analysis is another venue for channel geometry estimations. Hilldale and Raff (2007) evaluate LiDAR's effectiveness and find that the results display consistent bias. Kinzel et al. (2013) note that this form of LiDAR is still in the developmental/exploratory phase of usage. Legleiter et al. (2015) note that the bathymetric LiDAR utilized in their study was unable to detect shallow depths and accuracy was highly reduced in turbid streams. However, the prospect of using Space-Borne green-wavelength LiDAR to map bathymetry is promising though and will be of great interest in future studies (Adhallah et al., 2013). Using this LiDAR based process could result in much more accurate channel representations. Continued development in this field will be critical in riverine hydraulic model performance.

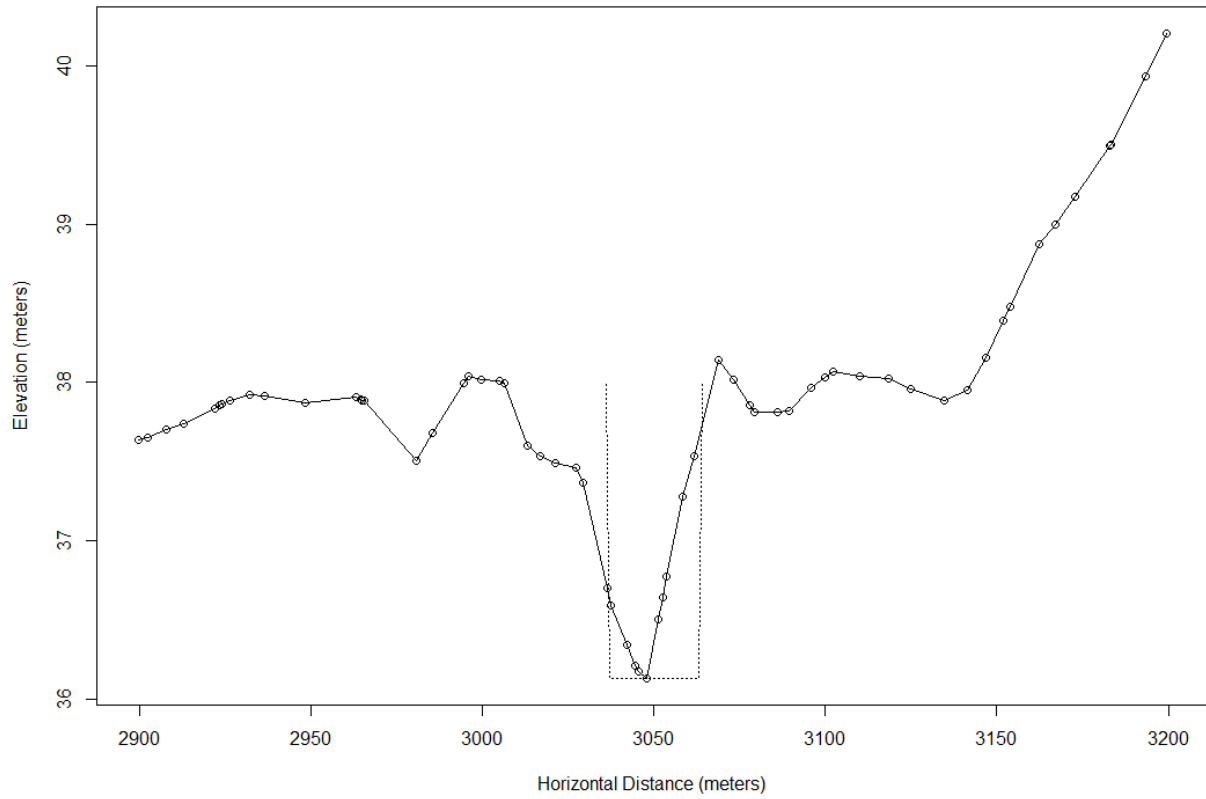


Figure 7. A cross section from the Blackwater River in Escambia County, Alabama. The solid line represents the surveyed cross-section with overbanks. The dotted line represents a rectangular hydraulic geometry estimate of the channel.

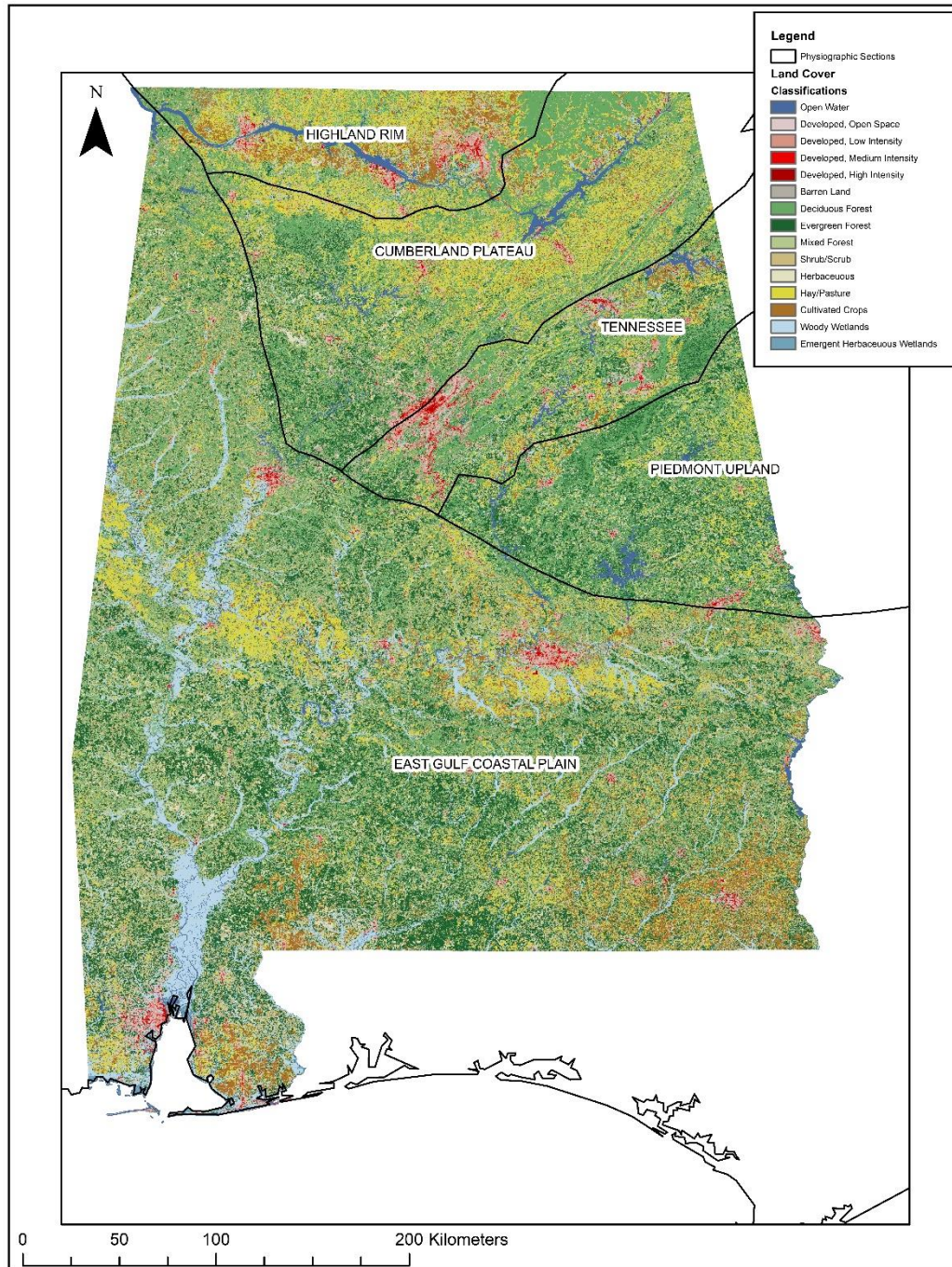


Figure 8. Land cover in Alabama. Note that the pervasive land cover in areas where hydraulic geometries do not match the HEC-RAS results is either developed with either agricultural or developed uses. The areas include the Highland Rim, Cumberland Plateau, and Tennessee physiographic sections.

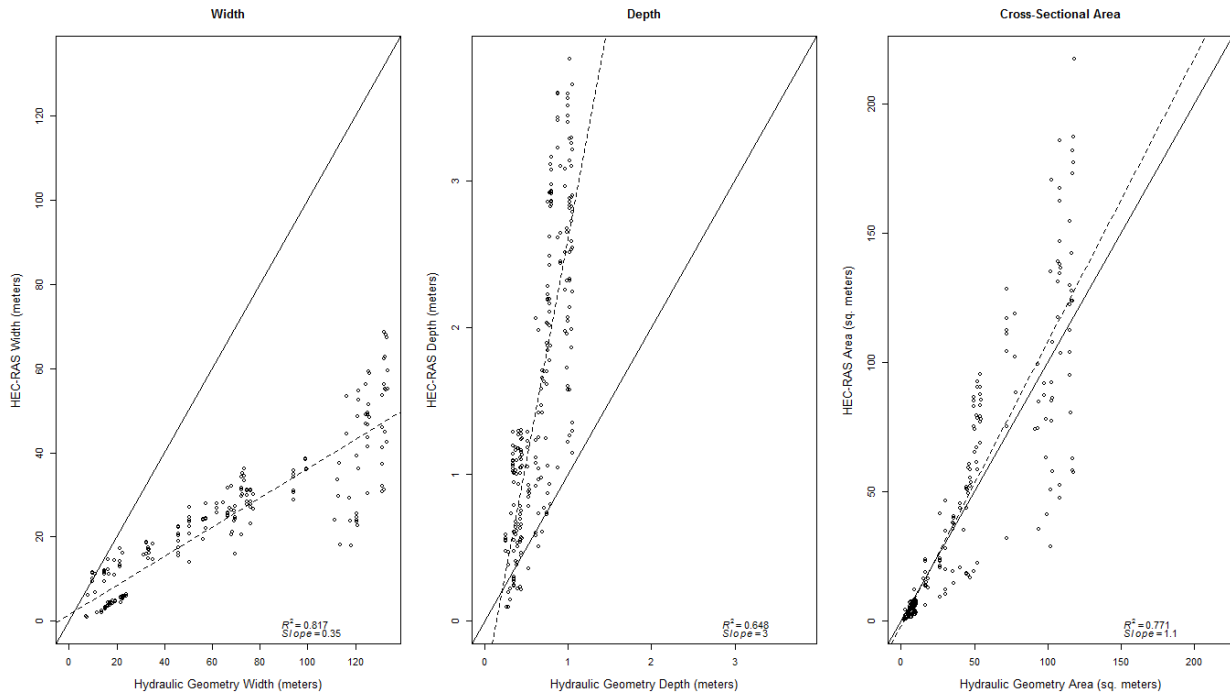


Figure 9. Flipped dimensions at ten times the naturalized flow in the Cumberland Plateau physiographic section of Alabama. Here the hydraulic geometry width is larger than those estimated by HEC-RAS and the depth is smaller than those estimated by HEC-RAS. When combined, the area of the channel approximation produced by the hydraulic geometries is approximate to that produced by HEC-RAS. The solid line represents a slope of one and the dashed line is the regression value. This result suggests that the estimates of width and depth for the Cumberland Plateau section of Alabama are opposite to those observed in other parts of the physiographic province.

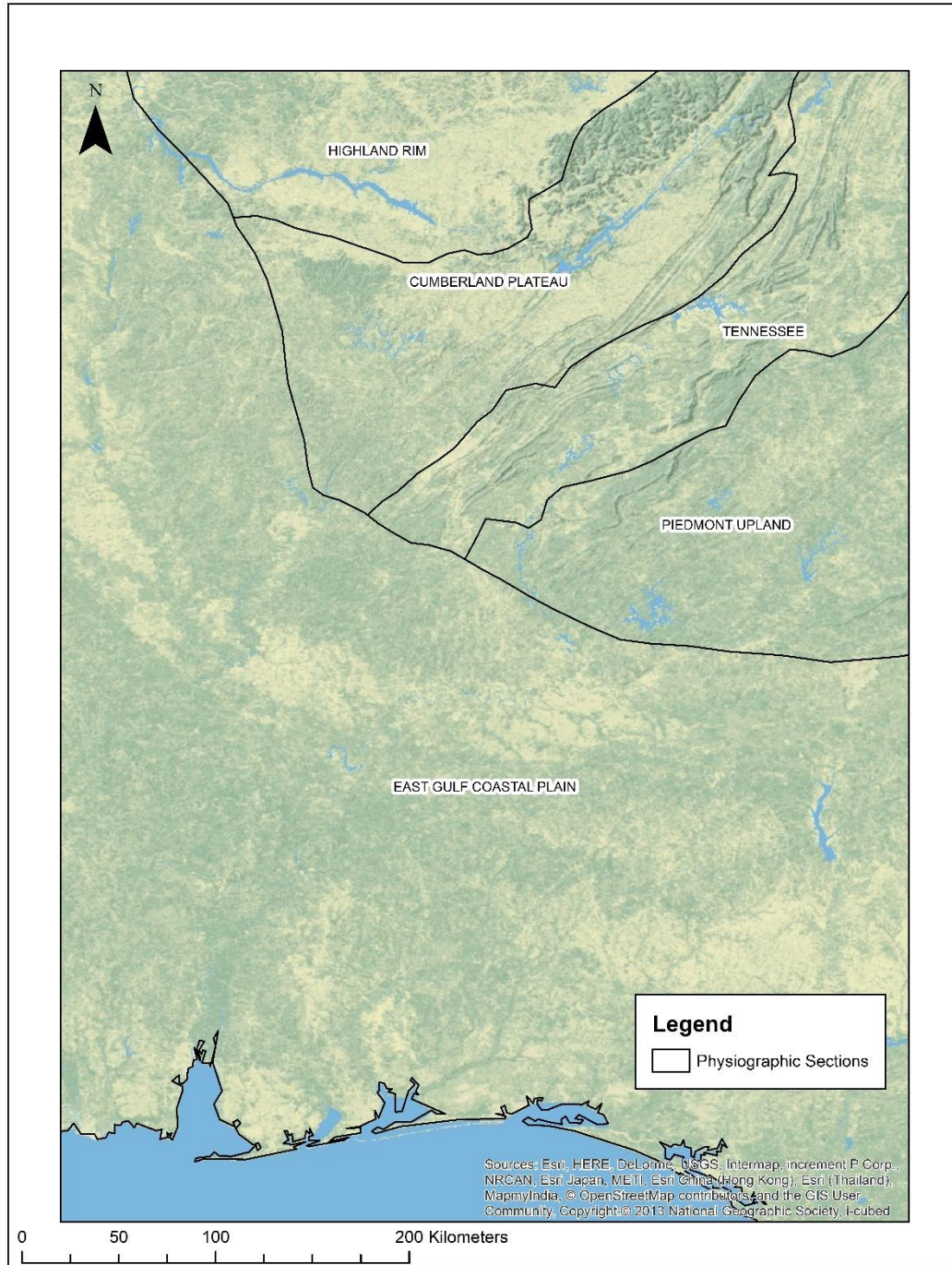


Figure 10. Terrain of Alabama. Note the relatively low gradient, homogeneity of the East Gulf Coastal Plain physiographic section as compared to the other physiographic sections of Alabama. This indicates that the stronger ability to predict channel geometry in this region maybe due to either the lack of relief or homogeneous relief within the relief.

Depth estimation

Correlation between HEC-RAS estimates of depth and hydraulic geometry depths was statistically significant in the East Gulf Coastal Plain and, to lesser degree, the Piedmont Upland physiographic regions. These results suggest that for a majority of the state of Alabama, hydraulic geometries are an accurate means of estimating water depth, given a discharge, in unregulated streams. These findings have a number of implications. First, the geometries are suitable for developing synthetic depth-discharge rating curves. Rating curves describe the flow-depth relationship are commonly used in modeling applications (Mueses et al., 2007). Second, the discharge-flow power relationships themselves may be used to calculate a reasonable approximation of water inundation, given an accurate DEM of the river and floodplain. Interestingly, this research posits that the use of hydraulic geometries may provide a depth-discharge relationship that works under any flow regime, extending their usage potential. These results correspond to the same physiographic sections where hydraulic geometries provide reasonable channel dimension approximations. The correlation between these results and those presented in the previous section suggests that depth estimation is subject to the same constraints as these previous findings on estimating stream channel dimensions.

The results within the East Gulf Coastal Plain and Piedmont Upland physiographic regions indicate that the geometries provide a more consistent estimate of depth at higher flows than at lower flow regimes. This correlation between high flows and more accurate channel depths is intriguing. The intended usage of hydraulic geometries is for bankfull discharge. Bankfull discharge roughly correlates to 1.5-year recurrence interval. We found that the magnitude of 100 times the gage adjusted naturalized flow is in many cases, greater than a 100 year event through comparison of the HEC-RAS model 100-year flows and the gage adjusted naturalized flow. In

such flow regimes, resulting R^2 from regressing the HEC-RAS depths on the hydraulic geometry depths were higher and slope terms from each regression were consistent with lower flows. Higher R^2 likely correspond to higher depths, at least partially, because of large flows producing consistently higher depths for both the HEC-RAS simulations and the hydraulic geometries.

Figure 11 compares the resulting hydraulic depths at each cross section from 100-Year Events in the East Gulf Coastal Plain dataset to the hydraulic geometries. The HEC-RAS depths represented within Figure 11 are hydraulic depths from 100-Year Events found using either USGS regional regression equations or from hydrologic models, such as the U.S. Army Corps of Engineers Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS). Here, an adequate R^2 is found, along with a slope which was consistent with all other multiples of gage adjusted naturalized flow. These results suggest that not only can hydraulic geometries measure flood depths but that they may be a more reliable under high flow conditions. This result is key in that the usage of these relationships is possible during flood events.

Water depth is the most common mechanism for estimating damage that occurs as a result of flooding. The Hazard United States Multi-Hazard (Hazardus) (FEMA, 2012), Hydrologic Engineering Center – Flood Impact Analysis (HEC-FIA) (USACE, 2012), and Hydrologic Engineering Center – Flood Damage Analysis (HEC-FDA) (USACE, 2008) are commonly used software in the United States to assess flood risk through depth-damage relationships. In the literature, the trend of depth-damage curve usage pervades over the course of several decades (Eckstein, 1965; Grigg and Helweg, 1975; Das and Lee, 1988; Yang and Tsai, 2000; Freni et al., 2010). The ability to reliably provide georeferenced water depth during floods is critical in understanding the damage floods cause. This research posits that a simple power function

relationship may capture this relationship without the use of sophisticated hydraulic or hydrodynamic modeling.

A speculative explanation for the goodness of fit between hydraulic geometries and HEC-RAS predicted water depths in the East Gulf Coastal Plain and Piedmont Upland for multiple flow regimes is that the power functions in these regions are more sensitive to the relationship between water depth and water discharge. Figure 12 shows hydraulic geometries as a function of increasing discharge for all 5 regions. For high discharge, note that some physiographic regions, such as the Cumberland Plateau, are very insensitive to changes in flow, as the streams in this region are relatively shallow. Other functions, such as those for the Piedmont Upland and East Gulf Coastal Plain, remain sensitive to changes in discharge at high flows, as the streams in this physiographic section are relatively deep. This phenomena may produce hydraulic geometries capable of estimating stream depth below and above a region's bankfull discharge. This observation requires additional study in other sections of the country to determine its validity.

Figure 14 demonstrates that in the East Gulf Coastal Plain, comparison between changes in water depth predicted by hydraulic geometries are roughly approximate to changes in water depth observed in USGS rating curves from the region (USGS, 2016). Figure 14 demonstrates that as discharge increases by $0.283 \text{ m}^3/\text{s}$ ($10 \text{ ft}^3/\text{s}$) the hydraulic geometries approximate USGS rating depth increases as gage sites from the same physiographic section. This is encouraging and supports the notion that hydraulic geometries in this region may provide a means to roughly approximate rating curves where none exist. Here, it appears as though the hydraulic geometries approximate the channels depth response to incremental changes in discharge. Additional research will analyze this finding to see if additional regions of the world might also demonstrate the same relationship.

This research asserts that first order estimates of flow depth are reasonable using hydraulic geometry coefficients for unregulated streams in areas where water depth is sensitive to discharge. If the results can be consistently proven in additional regions, these simple power function relationships would reduce the need for cross-sectional information and mapping flooding may only require knowledge of average streamflow (e.g. from the NHDPlus Version 2 flowlines) and an accurate DEM. We do not assert that hydraulic geometry relationships are a replacement for physical hydraulic modeling. Rather, possible usage scenarios must be determined by the user. These use cases are yet to be determined, but must be capable of absorbing significant uncertainty.

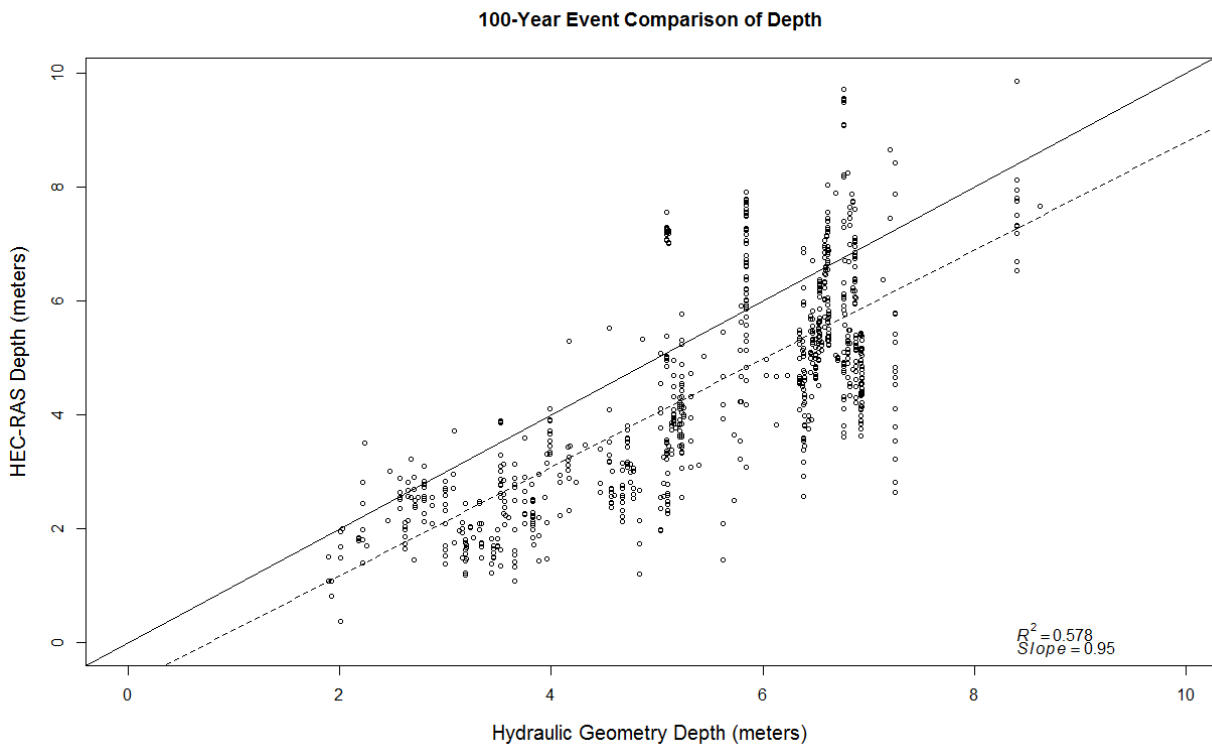


Figure 11. Comparison of estimated depths generated by hydraulic geometries and HEC-RAS for 100-Year events in the East Gulf Coastal Plain. The results are consistent with the output generated using all multiples of gage adjusted naturalized flow from the NHDPlus Version 2 dataset. The solid line represents a slope of one and the dashed line is the regression line.

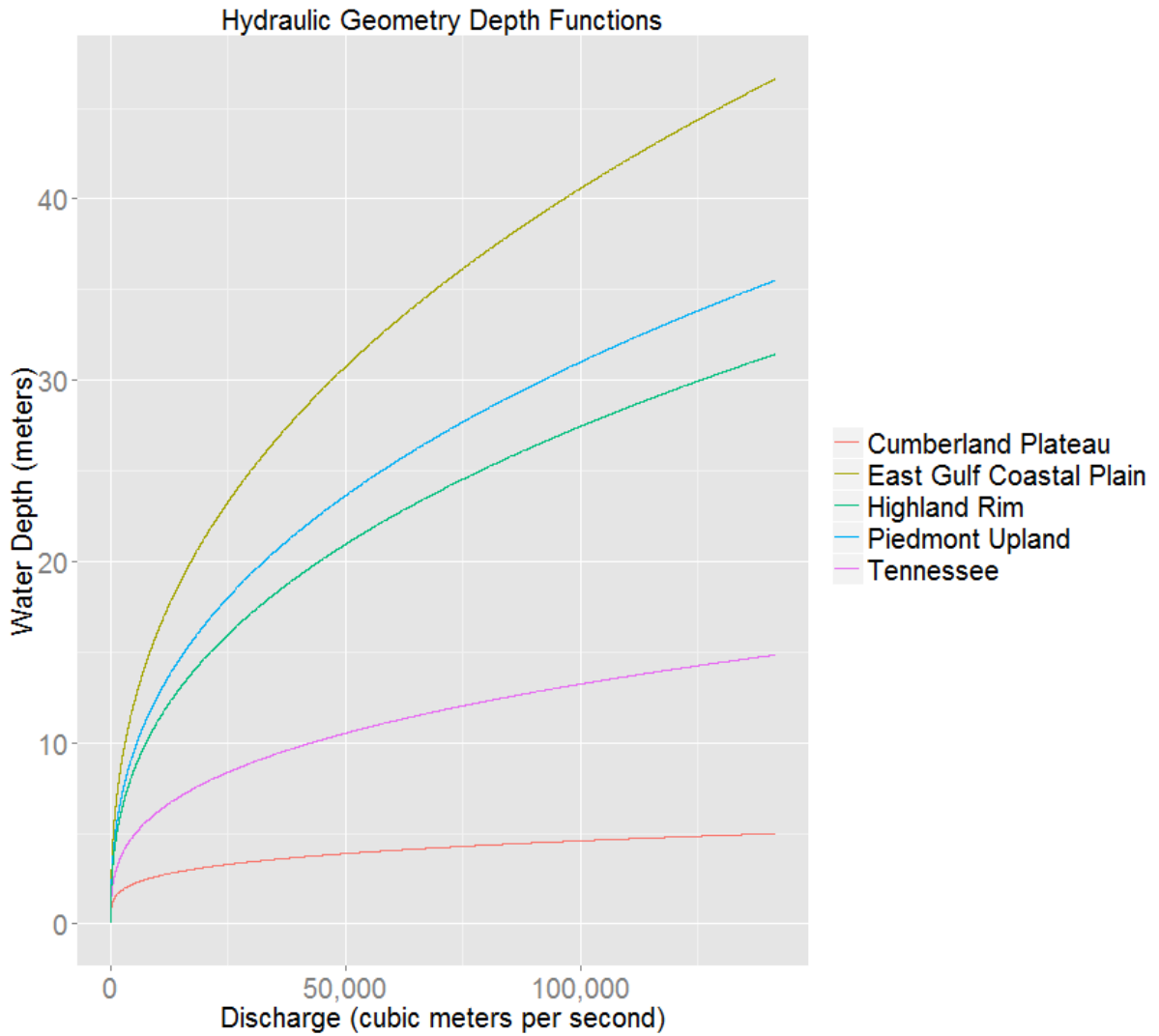


Figure 12. Hydraulic geometry depth functions used for each physiographic section in Alabama. Note that the functions which allow for accurate depth predictions at multiple flow regimes are those which possess a relatively high sensitivity to discharge. This characteristic may be the cause for the functions being able to accurately measure depth at relatively low and high flow regimes.

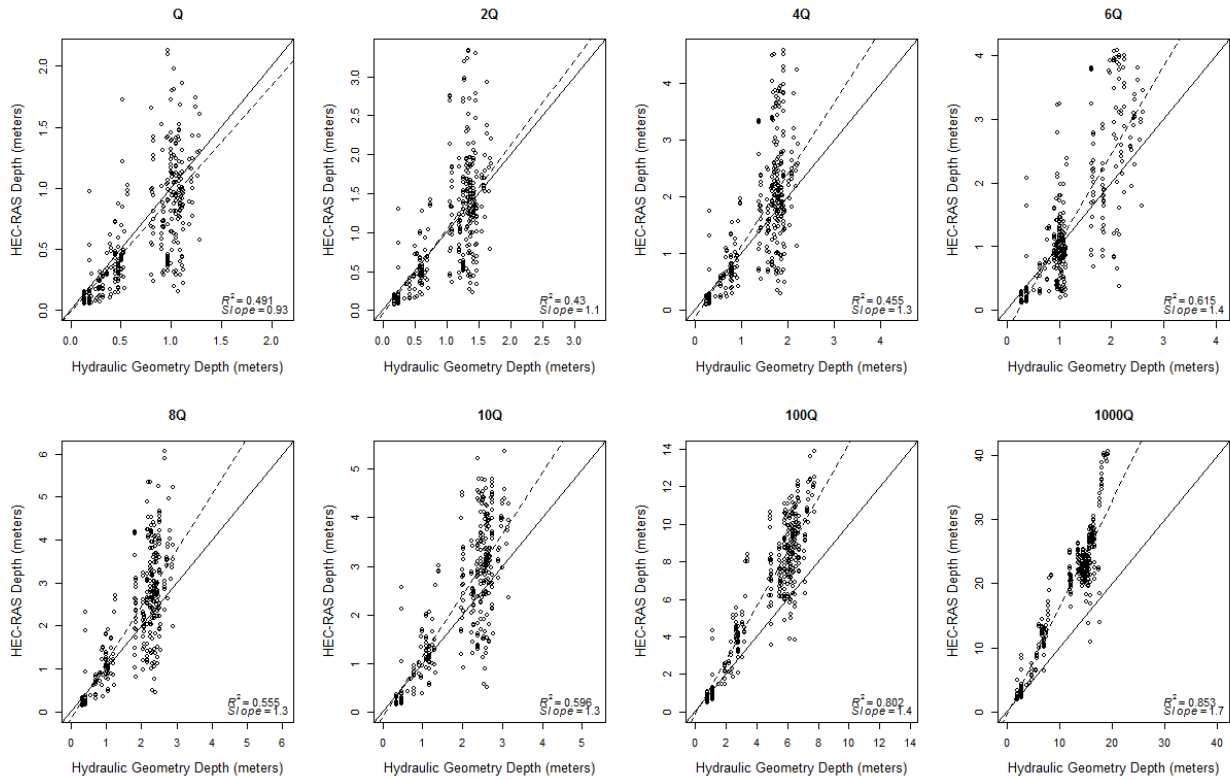


Figure 13. Relationship between HEC-RAS predicted depth and hydraulic geometry predicted depth for each multiple of gage adjusted naturalized flow (Q) for all Piedmont Upland sites (Figure 2). Note that the R^2 and slope are not as significant as in the East Gulf Coastal Plain but the slope is still very consistent throughout all flow regimes. The solid line represents a slope of one and the dashed line is the regression line.

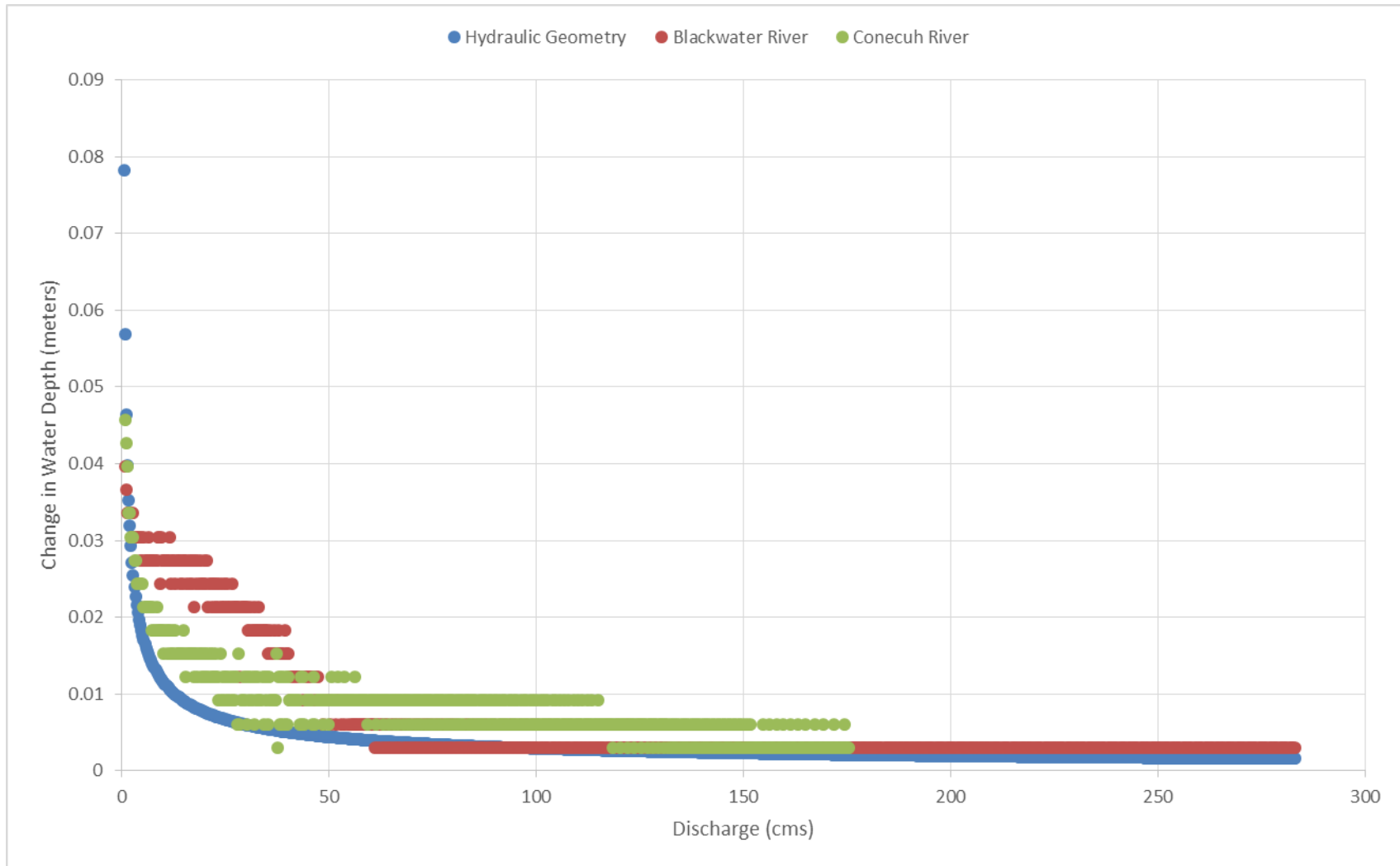


Figure 14. Comparison of Predicted Changes in Water Depth Based on Hydraulic Geometries and Rating Curves for USGS gage locations along the Blackwater and Conecuh Rivers. These rivers are in the East Gulf Coastal Plain.

References

- Abdallah, H., J. Bailly, N. N. Baghdadi, N. Saint-Geours, and F. Fabre. (2013). Potential of space- borne LiDAR sensors for global bathymetry in coastal and inland waters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(1), 202-216. DOI: 10.1109/JSTARS.2012.2209864
- Archfield, S. A., M. Clark, B. Arheimer, L. E. Hay, H. McMillan, J. E. Kiang, et al. (2015). Accelerating advances in continental domain hydrologic modeling. *Water Resources Research*. DOI: 10.1002/2015WR017498
- Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. (1998). Large area hydrologic modeling and assessment part I: Model development. *Journal of the American Water Resources Association (JAWRA)*, 34(1), 73-89. DOI: 10.1111/j.1752-1688.1998.tb05961.x
- Bent, G. C., & Waite, A. M. (2013). Equations for Estimating Bankfull Channel Geometry and Discharge for Streams in Massachusetts. (No. 2013-5155). Reston, Virginia: US Geological Survey.
Retrieved from <http://pubs.usgs.gov/sir/2013/5155/pdf/sir2013-5155.pdf>
- Bicknell, B. R., Imhoff, J. C., Kittle Jr, J. L., Donigan Jr, A. S., & Johanson, R. C. (1997). Hydrological simulation program--fortran: User's manual for version 11. US Environmental Protection Agency, National Exposure Research Laboratory Athens, GA.
- Bieger, K., H. Rathjens, P. M. Allen, and J. G. Arnold. (2015). Development and Evaluation of Bankfull Hydraulic Geometry Relationships for the Physiographic Regions of the United States. *Journal of the American Water Resources Association (JAWRA)*, 51(3), 1-17, DOI: 10.1111/jawr.12282.
- Bondelid, T. (2014). EROM monthly flows. Horizon Systems.
- Brockman, R. R. (2010). Hydraulic Geometry Relationships and Regional Curves for the Inner and Outer Bluegrass Regions of Kentucky. Master's Thesis, University of Kentucky, Lexington, Kentucky.
- Brunner, G. W. (2010). HEC-RAS river analysis system user's manual version 4.1. (Computer Program Documentation No. CPD-68). Davis, CA: U.S. Army Corps of Engineers.
- Castro, J. M. and P. L. Jackson. (2001). Bankfull Discharge Recurrence Intervals and Regional Hydraulic Geometry Relationships: Patterns in the Pacific Northwest, USA. *Journal of the American Water Resources Association (JAWRA)*, 37(5), 1249-1262.
- Das, S. and R. Lee. (1988). A Nontraditional Methodology for Flood Stage-Damage Calculations. *Journal of the American Water Resources Association (JAWRA)*, 24(6), 1263-1272.

- Davies, J. R., K. M. Lagueux, B. Sanderson, and T. J. Beechie. (2007). Modeling Stream Channel Characteristics From Drainage-Enforced DEMs in Puget Sound, Washington, USA. *Journal of the American Water Resources Association (JAWRA)*, 43(2), 414-426. DOI: 10.1111/j.1752-1688.2007.00032.x
- Eckstein, O. (1965). *Water Resource Development – The Economics of Project Evaluation*. Harvard University Press, Cambridge, Massachusetts.
- Federal Emergency Management Agency (FEMA). (2012b). Hazus-MH user manual. (Computer Program Documentation). Washington, D.C.: Federal Emergency Management Agency.
- Fenneman, N. M. (1917). Physiographic Subdivision of the United States. *Proceedings of the National Academy of Sciences of the United States of America*, 3(1), 17-22.
- Freeman, P., and S. Northcutt. (2014). Guess Creek Watershed. *25th Annual Nonpoint Source Conference*. Alabama Department of Environmental Management (ADEM), Montgomery, Alabama.
- Freni, G., G. La Loggia, and V. Notaro. (2010). Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation. *Water Science and Technology*, 61(12), 2979-2993. DOI: 10.2166/wst.2010.177
- Grigg, N. S. and O. J. Helweg. (1975). State-of-the-Art of Estimating Flood Damage in Urban Areas. *Journal of the American Water Resources Association (JAWRA)*, 11(2), 379-390.
- Hardison, E. C., O’Driscoll, M. A., DeLoatch, J. P., Howard, R. J., and M. M. Brinson. (2009). Urban land use, channel incision, and water table decline along coastal plain streams, North Carolina. *Journal of the American Water Resources Association (JAWRA)*, 45(4), 1032-1046. DOI: 10.1111/j.1752-1688.2009.00345.x
- Hilldale, R. C. and D. Raff. (2008). Assessing the ability of airborne LiDAR to map river bathymetry. *Earth Surface Processes and Landforms*, 33(5), 773-783. DOI: 10.1002/esp.1575
- Hodges, B. R. (2013). Challenges in continental river dynamics. *Environmental Modelling & Software*, 50, 16-20. DOI: 10.1016/j.envsoft.2013.08.010
- Horritt, M. and P. Bates. (2002). Evaluation of 1D and 2D numerical models for predicting river flood inundation. *Journal of Hydrology*, 268(1), 87-99. DOI: 10.1016/S0022-1694(02)00121-X
- Hupp, C. R. (2000). Hydrology, geomorphology and vegetation of Coastal Plain rivers in the south-eastern USA. *Hydrological Processes*, 14(16-17), 2991-3010. DOI: 10.1002/1099-1085(200011/12)14:16/173.0.CO;2-H

- Johnson, P., and Fecko, B. (2008). Regional channel geometry equations: a statistical comparison for physiographic provinces in the eastern US. *River Research and Applications*, 24(6), 823-834. DOI: 10.1002/rra.1080
- Kinzel, P. J., C. J. Legleiter, and J. M. Nelson. (2013). Mapping River Bathymetry With a Small Footprint Green LiDAR: Applications and Challenges1. *Journal of the American Water Resources Association* (JAWRA), 49(1), 183-204. DOI: 10.1111/jawr.12008
- Legleiter, C. J., B. T. Overstreet, C. L. Glennie, Z. Pan, J. C. Fernandez-Diaz, and A. Singhanian. (2015). Evaluating the capabilities of the CASI hyperspectral imaging system and Aquarius bathymetric LiDAR for measuring channel morphology in two distinct river environments. *Earth Surface Processes and Landforms*. DOI: 10.1002/esp.3794
- Leopold, L. B. and T. Maddock Jr. (1953). The Hydraulic Geometry of Stream Channels and Some Physiographic Implications. U.S. Geological Survey Professional Paper 252, 57 pp. <http://pubs.er.usgs.gov/publication/pp252>
- Magirl, C. S. and T. D. Olsen. (2009). Navigability Potential of Washington Rivers and Streams Determined with Hydraulic Geometry and a Geographic Information System. U.S. Geological Survey Scientific Investigations Report 2009-5122, 21 pp. <http://pubs.usgs.gov/sir/2009/5122/pdf/sir20095122.pdf>
- Maidment, D. R. (2014). National Flood Interoperability Experiment. *2014 American Geophysical Union (AGU) Fall Meeting*. American Geophysical Union, San Francisco, California.
- McBride, M., and D. B. Booth. (2005). Urban Impacts on Physical Stream Condition: Effects of Spatial Scale, Connectivity, and Longitudinal trends. *Journal of the American Water Resources Association* (JAWRA), 41(3), 565–580. DOI: 10.1111/j.1752-1688.2005.tb03755.x
- McCandless, T. L. (2003a). Maryland stream survey: Bankfull discharge and channel characteristics of streams in the coastal plain hydrologic region. (No. CBFO-S03-02). U.S. Fish & Wildlife Service. Retrieved from <http://www.fws.gov/chesapeakebay/pdf/plain.pdf>
- McCandless, T. L. (2003b). Maryland stream survey: Bankfull discharge and channel characteristics of streams in the allegheny plateau and the valley and ridge hydrologic regions. (No. CBFO-S03-01). Annapolis, MD: U.S. Fish and Wildlife Service.
- Mersel, M. K., L. C. Smith, K. M. Andreadis, and M. T. Durand. (2013). Estimation of river depth from remotely sensed hydraulic relationships. *Water Resources Research*, 49(6), 3165-3179. DOI: 10.1002/wrcr.20176

- Merwade, V., A. Cook, and J. Coonrod. (2008). GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. *Environmental Modelling & Software*, 23(10), 1300-1311. DOI: 10.1016/j.envsoft.2008.03.005
- Merwade, V. (2014). Research activities. Retrieved from <http://web.ics.purdue.edu/~vmerwade/research.html>
- Mohamoud, Y. M. and R. S. Parmar. (2006). Estimating Streamflow and Associated Hydraulic Geometry, The Mid-Atlantic Region, USA. *Journal of the American Water Resources Association* (JAWRA), 42(3), 755-768. DOI: 10.1111/j.1752-1688.2006.tb04490.x
- Mueses, A., A. Said, and M. Ross. (2007). Generalized Nondimensional Depth-Discharge Rating Curves Tested on Florida Streamflow. *Journal of the American Water Resources Association* (JAWRA), 43(2), 473-481. DOI: 10.1111/j.1752-1688.2007.00037.x
- Natural Resource Conservation Service (NRCS). Regime curve for rural piedmont plateau physiographic province. Retrieved from http://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/water/quality/?cid=nrcs143_015053
- Oubeidillah, A. A., S. Kao, M. Ashfaq, B. S. Naz, and G. Tootle. (2014). A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US. *Hydrology and Earth System Sciences*, 18(1), 67-84. DOI: 10.5194/hess-18-67-2014
- Raymond, D. E., W. E. Osborne, C. W. Copeland, and T. L. Neathery. (1988). Alabama Stratigraphy. <http://www.gsa.state.al.us/documents/pubs/onlinepubs/Circulars/C140.PDF>
- Rosgen, D. L. (1994). A classification of natural rivers. *Catena*, 22(3), 169-199.
- Smoot, Z., A. Jayakaran, and D. Hitchcock. (2014). Hydraulic Geometry Curves and Bankfull Recurrence in the Pee Dee River Basin. *Proceedings of the 2014 South Carolina Water Resources Conference*. Clemson University, Columbia, South Carolina.
- Sood, A. and V. Smakhtin. (2015). Global hydrological models: a review. *Hydrological Sciences Journal*, 60(4), 549-565. DOI: 10.1080/02626667.2014.950580
- Stewardson, M. (2005). Hydraulic geometry of stream reaches. *Journal of Hydrology*, 306(1), 97-111. DOI: 10.1016/j.jhydrol.2004.09.004
- U.S. Army Corps of Engineers (USACE). (2008). HEC-FDA flood damage reduction analysis user's manual. (Computer Program Documentation No. CPD-72). Davis, CA: U.S. Army Corps of Engineers (USACE).

- U.S. Army Corps of Engineers (USACE). (2012). HEC-FIA flood impact analysis: User's manual. (Computer Program Documentation No. CPD-81). Davis, CA: U.S. Army Corps of Engineers.
- United States Geological Survey (USGS). (2013). Mean daily flow for unregulated streams. Retrieved from <http://al.water.usgs.gov/duration.plots/unregulated.html>
- United States Geological Survey (USGS). (2015). The national map: Elevation. Retrieved from <http://nationalmap.gov/elevation.html>
- United States Geological Survey (USGS). (2016). Waterwatch. Retrieved from <http://waterwatch.usgs.gov/index.php>
- Walter, R. C. and D. J. Merritts. (2008). Natural streams and the legacy of water-powered mills. *Science*, 319(5861), 299-304. DOI: 10.1126/science.1151716
- Yang, C. and C. Tsai. (2000). Development of a GIS-Based Flood Information System for Floodplain Modeling and Damage Calculation. *Journal of the American Water Resources Association (JAWRA)*, 36(3), 567-577. DOI: 10.1111/j.1752-1688.2000.tb04287.x
- Yoon, Y., M. Durand, C. J. Merry, E. A. Clark, K. M. Andreadis, and D. E. Alsdorf. (2012). Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of hydrology*, 464, 363-375. DOI: 10.1016/j.jhydrol.2012.07.028

CHAPTER 3 RAPID FLOOD DAMAGE ASSESSMENT USING PUBLIC DOMAIN CADASTRAL AND ADDRESS POINT DATA WITH FUZZY LOGIC ALGORITHMS

Introduction

The National Flood Interoperability Experiment (NFIE) sought to demonstrate the capability of improved flood predictive capacity of the conterminous United States (CONUS). NFIE-derived technologies and workflows offer the ability to rapidly forecast flood damages that coincide with the hydrologic and hydraulic estimations physics-based models generate. NFIE-Geo, a geodatabase composed of data requisite for large scale hydrologic modeling, will expand to provide access to a CONUS-wide set of Address Points for emergency management. Most U.S. County tax assessment offices throughout CONUS produce georeferenced cadastral data. To varying degrees, these parcel data describe building characteristics of structures within the parcel. Joining Address Point data with cadastral data offers the ability to rapidly develop structures inventories similar to that which the Federal Emergency Management Agency (FEMA) United States Hazard – Multi Hazard (Hazus) utilizes. In this work, the authors evaluate a purposed methodology which relies upon Address Point and parcel datasets to forecast damage estimates to buildings in CONUS. Results indicate that using Address Point and cadastral datasets can generate realistic flood damage estimations. The effort highlights an initial means by which the functionality of Hazus can change in future iterations. A case is made for the inclusion of cadastral data within the Open Water Data Initiative (OWDI).

Review of Literature

Traditionally, flood damage assessment occurs in a desktop computing environment with a mélange of meteorological, hydrologic, hydraulic, and damage assessment tools (Apel et al, 2004). In the United States, a range of these flood damage assessment softwares is available to local, state, and federal emergency management agencies. Among the more commonly used and freely accessible models are the Hydrologic Engineering Center's Flood Damage Reduction Analysis (HEC-FDA), Flood Impact Assessment (HEC-FIA), and FEMA's Hazus. Generally, each of these models assess direct damage to residential, commercial, agricultural, and industrial structures based upon the depth of floods and associated damage at those depths (USACE, 2008; USACE, 2012; FEMA, 2012a; FEMA, 2012b).

Each of these damage assessment models focuses on empirically and synthetically derived depth-damage relationships that describe proportionate damage for a structure under certain water depths. Depth-damage functions are the flood damage assessment mechanism most commonly used in the literature (Eckstein, 1965; Grigg and Helweg, 1975; Das and Lee, 1988; Yang and Tsai, 2000; Freni et al., 2010). The depth-damage curves themselves come from a list of sources, including the U.S. Army Corps of Engineers (USACE) (USACE, 2003; Davis and Skaggs, 1992) and FEMA (NFIP, 2013). However, each of the three models has its own unique methodology of assessing the relationship between flood depth and damage. HEC-FDA assesses direct damage through Monte-Carlo simulation of hydrographs, generating a distribution of peak flows. These peak flows produce a distribution of flood depths, each of which has an associated annualized damage calculation (USACE, 2008). HEC-FIA assesses direct damage by instead utilizing a single hydrograph. For each time step on the hydrograph, assessment of damage functions occurs. This methodology is ideal for simulation of single

flood event impacts (USACE, 2012). Hazus can assess direct damage through frequency-related analysis, providing annualized losses, or can analyze the damage associated with a single event or specific discharge (FEMA, 2012b).

Studies conducted by Banks et al. (2014) and Ding et al. (2008) evaluate the functionality of HEC-FDA, HEC-FIA, and Hazus. Banks et al. (2014) performed a comprehensive review of available software's geared toward flood planning and found that Hazus was the only one of the 11 reviewed to meet all of their evaluation criteria. This review included HEC-FIA, though Banks et al. (2014) did not include HEC-FDA in their review. Ding et al. (2008) assessed Hazus' performance at both Level I and Level II performance and compared model construction and outputs against a previously conducted study that utilized HEC-FDA. Hazus Level I analyses use only prepackaged datasets and functionality within Hazus. Hazus Level II analyses use higher-resolution data and third-party software. The results of Ding et al. indicate that the outputs of the Level II Hazus simulation were very similar to those created in HEC-FDA, but required a fraction of the man-power used to develop the latter.

Hazus has several limitations including a dependency on a single software suite, no applied programming interface (API), and a need for regional datasets that are difficult to develop and procure. Hazus functions within one combination of one proprietary software and operating system. An additional lack of an API precludes the user from running scripted analyses, such as Monte Carlo simulation. For example, Tate et al. (2014) had to utilize third party software in their Monte Carlo simulation experiment. Further, the developers of Hazus indicate that general datasets that they produce are viable only as baseline analyses. Regional Level II datasets are necessary for accurate damage analysis (Ding et al., 2008; Tate et al., 2015; FEMA, 2012b).

However, computational hydrology and hydraulics are moving into the realm of continental scale simulation and forecasting in parallel computing environments. The National Flood Interoperability Experiment (NFIE) held at the National Water Center (NWC) in Tuscaloosa, Alabama led to the development of a large scale stream flow forecasting framework capable of 15-hour lead-time forecasts for the entire coterminous United States (CONUS) (Maidment, 2014). This national water model will become operational in 2016 at the NWC (Gochis, 2015). Further, a parallel framework estimating inundation is under development. This large-scale application of the water cycle presents an opportunity to expand the capabilities of flood damage assessment.

A break from the traditional Hazus model exists within the Hazus R package that provides the depth-damage functions utilized within Hazus. The Hazus R package provides the first instance of modularizing the inner mechanics of Hazus. The R package gives a user access to approximately 1,300 depth-damage relationships stored within the Hazus software's flood module (Goteti, 2014). Thus, while only a small portion of the Hazus functionality, the Hazus R package provides a glimpse into the future of flood damage assessment.

Flood damage prediction can provide a range of benefits: 1) a fundamental first step in understanding economic impact of a flood event; 2) a mechanism for flood risk mapping; 3) an ability to proactively decrease economic impact to floods and; 4) a means to more effectively allocate emergency funds prior to flooding events (Oliveri and Santoro, 2000; Merz et al., 2010; Pappenberger et al., 2015). Spatial data containing requisite economic profiles for structure damage evaluation are available nationwide through each U.S. county's Tax Assessor department. Collectively, the Committee on Land Parcel Databases (2007) call these datasets cadastral data. Cadastral data is a primary input of most Hazus models (Ding et al., 2008; Tate et al.,

2014). Though disparate in structure, stored in various locations, and maintained by each individual county (Pita et al. 2011; FEMA 2012b), these public data hold significant value when paired with Address Point data for every street address in the United States. Address Point data also vary in composition. Address Point data can approximate building locations within the parcel. Combined, the cadastre and Address Point data sets can form Hazus inputs. Further, Address Point and cadastre data are routinely updated, unlike traditional Level II Hazus datasets. Therefore, they construct dynamic datasets for inclusion into a real-time, flood damage evaluation tool.

Approximately 70% of these cadastre datasets are digital, as of 2007 (Committee on Land Parcel Databases 2007). Cadastral are available online in most states, such as those in Kentucky (Kentucky Property Valuation Administrators, 2006), Alabama (Flagship GIS, Inc.), and Texas (OnlineSearches, 2015). Further, Google Earth Pro includes a subset of cadastre data in the United States (Google, 2015). The company ReportAll also maintains a database of parcel information that covers approximately 92.4% of the United States' population (ReportAll, 2015).

Though the Hazus-MH User's Manual lists these datasets as a primary means of constructing Level II regional data (FEMA, 2012b), their existing application in large-scale scenarios is not pervasive. Bennett et al. (2010) note their importance in natural hazards research, particularly in preparation, response, and recovery. The Committee on Land Parcel Databases (2007) also highlights these data's importance. Sperling and von Meyer (2013) describe an exploratory project commissioned by the U.S. Department of Housing and Urban Development (HUD) which seeks to standardize a national database of parcel information composed at the county tax assessor offices across the country. Sperling and von Meyer (2013)

conclude that such a dataset is technically feasible, though many obstacles remain before this dataset can be realized (Abt Associates, Inc. 2013).

Working under the assumption that a variable structure to the cadastre datasets is certain in the near term future, programs making use of the data must adapt to this situation. Key to correctly modeling flood damage is identifying the appropriate depth-damage relationship, as even perfect matches between a structure and depth-damage relationship do not account for all components of variability (Zhai et al., 2005; Tate et al., 2014). Cadastral data can provide a means of describing the building(s) at a particular location. Fuzzy logic interpretation using machines may provide a means to match approximate text descriptions of the cadastre and depth-damage functions.

Zadeh (1965) introduced the concept of fuzzy sets as a means to represent the imprecise nature of real world events in computer science. In general, fuzzy sets take the form

$$\mu_A(x) \in [0,1] A = (x, \mu_A(x) | x \in X) \quad (2)$$

Where A is a subset of the universe of discourse, X , $\mu_A(x)$ is a membership function which defines the degree to which element x is a member of subset A . As $\mu_A(x)$ approaches 1, it more closely resembles the subset A and as it moves towards 0, it is less like the subset A (Bai et al., 2007). Fuzzy logic follows fuzzy set theory and identifies the approximate nature of many-valued relationships (Novák, 2012).

Utilization of fuzzy logic occurs in numerous water resource investigations. These investigations include the identification of flood vulnerability (Lee et al. 2015) and stream waste-load allocation (Chen and Chang 2006). In both contexts, the researchers used fuzzy logic as an aid to multi-criteria decision making. The largest proportion of fuzzy logic applications occur as aids to decision making in water resources.

In the case of string matching, the Python module *FuzzyWuzzy* operates as a fuzzy logic algorithm. The company *SeatGeek* developed *FuzzyWuzzy* to identify event tickets from various sources (Cohen 2011). This module supports the ratio text matching process where, according to the Python website (Python Software Foundation, 2016)

$$R = \frac{2M}{T} \quad (3)$$

where the ratio metric (R) is equal to the quotient of twice the total number of matches in a string comparison (M) and the summed total number of elements in those two strings (T). Using Equation 3, a better match will have a value which approaches one. Thus, R is equivalent to $\mu_A(x)$, the membership function.

Method

Fuzzy logic was chosen because of extended usage in a number of scenarios, its ability to provide the desired outcome for this project, and its ease of implementation. Fuzzy logic is attune for problems involving imprecise outcomes (Zadeh, 1965). In this research, the problem the researchers look to investigate is how to best approximate a match between a building and depth-damage curve using each respective text description. These text descriptions are unlikely that to exactly match one another. Therefore, the task is to determine the best match from a set of many possible matches.

Fuzzy logic, in particular fuzzy text matching, are attune for such tasks involving imprecision. Fuzzy logic can classify and retrieve data when sporadic misspellings occur within a database (Rees, 2014). Fuzzy string matching also serves as a means of integrating similar records (Huang and Madey, 2004). In effect, fuzzy text matching is an adaptive and efficient means of classifying and merging data when an exact match between two strings is unlikely.

The fuzzy logic system implemented *FuzzyWuzzy*, requires minimal overhead to learn and set up and is an open source solution. The Python *FuzzyWuzzy* module (SeatGeek, 2015) provides the mechanism for matching a cadastre or parcel point dataset to the most appropriate depth-damage function by comparing the individual text descriptions of each dataset. *FuzzyWuzzy* adapts to the variability of each dataset. *FuzzyWuzzy* requires no training to utilize its algorithms. Using Equation 3, a better match will have a value which approaches one. Here the researchers use the *token_set_ratio* function within *FuzzyWuzzy*. The *token_set_ratio* function tokenizes each word in both text strings and finds where the two strings have identical words (tokens) or intersections. This grouping is then compared, using Equation 1, to remainders which are unique to each string. In this context, ratio scores are higher for string comparisons where overlaps are a larger proportion of the entire string or where the remainders are similar to the intersections (Cohen, 2011). As a rule of thumb, a ratio score of 0.6 (60) or greater is usually indicative of a good match between two strings (Python Software Foundation, 2006). The software omits those Address Point-Cadastre sites not meeting the ration score specification or those describing undeveloped or vacant locations.

Depth-damage and depreciation functions within the Hazus R package combine with a set of Python 2.7 scripts to form the tool *Flood Damage Wizard*. *Flood Damage Wizard* performs these tasks, which are also outlined in Figure 15:

1. It provides a set of tools for blending cadastre data with any type of point shapefile.
2. It extracts water depths at each Address Point-Cadastre location from a user-supplied inundation grid produced within a riverine hydraulic model.

3. It determines which depth-damage curve best fits the description of each Address Point-Cadastre location. Depth-damage curves exist to describe damage to the building's structure, content, and inventory.
4. It calculates financial and economic damage to each Address Point-Cadastre location using a combination of the matching depth-damage curve, Address Point-Cadastre data, and generalized data (Table 3).
5. It generates seven different scenarios for each inundated Address Point-Cadastre location:
 - 1) structural, content, and inventory damage
 - 2) structural and content damage
 - 3) structural and inventory damage
 - 4) content and inventory damage
 - 5) structural damage
 - 6) content damage
 - 7) inventory damage.

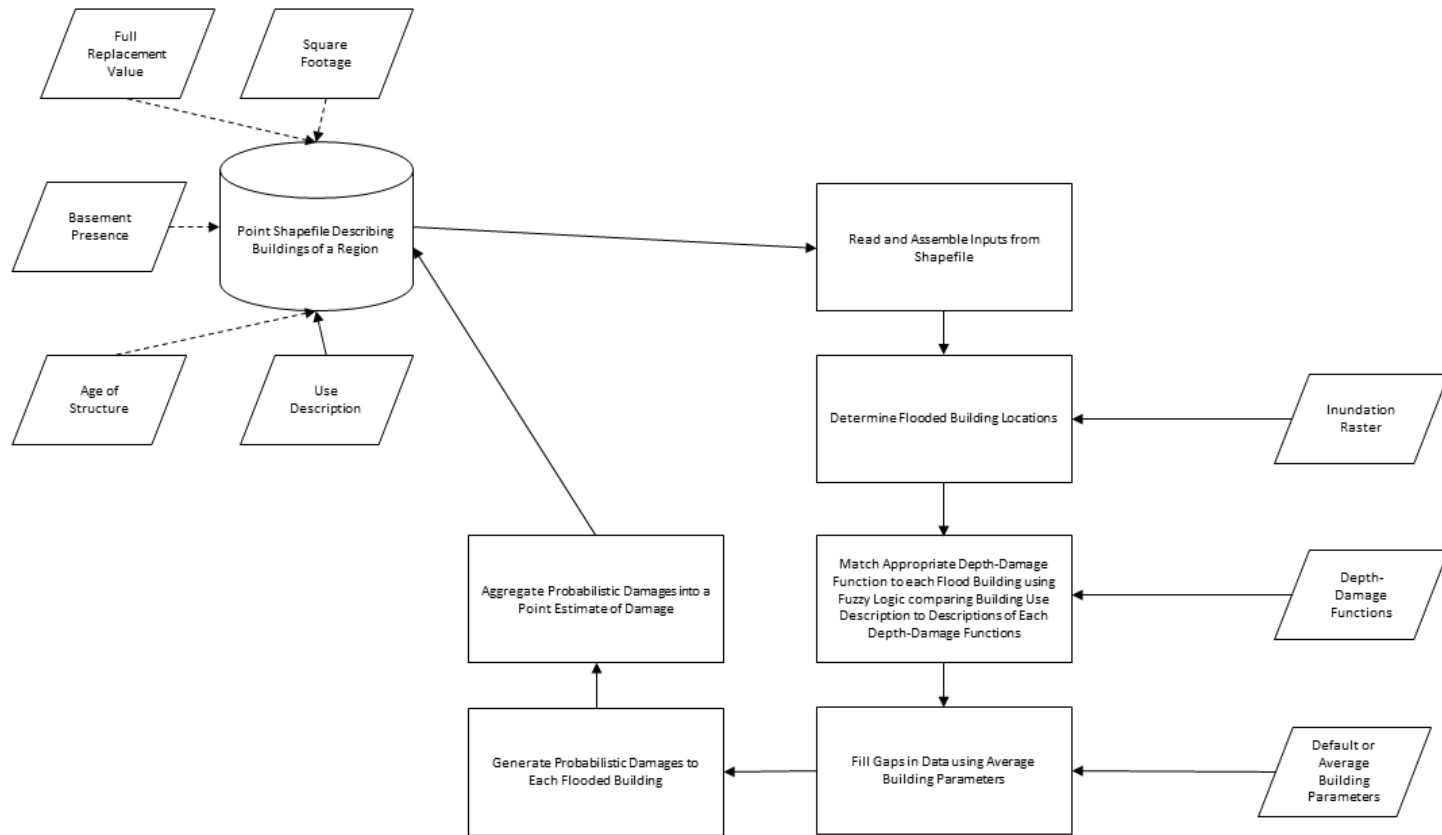


Figure 15. System flowchart for flood damage assessment software (Flood Damage Wizard).

As an example, if the parcel description “single family two story” cycles through the system, *Flood Damage Wizard* compares this text description to each text description in the library of depth-damage function text descriptions. *Flood Damage Wizard* determines the single, best matching curve description to be the one which generates the highest ratio score. In this instance, the best match to “single family two story” is “two story w/ 1/2 living area single family residential dwelling”. The depth-damage functions for the structure, content, and inventory which have the “two story w/ 1/2 living area single family residential dwelling” description calculate the flood damage or loss.

Chapter 14 of the Hazus Flood Technical Manual (2012) specifies a means of estimating structural, content, and inventory financial and economic damage during floods for given occupancy classes. Financial damages represent full replacement value of the structure, contents, and inventory. Economic damages represent depreciated full replacement value, which accounts for wear and tear to consumer items (Penning-Roswell et al., 2003; Penning-Roswell et al., 2005; Merz et al., 2010). Depreciated damage estimates are necessary for economic damage evaluation as full replacement values overestimate damage (Merz et al., 2010). Table 3 summarizes a list of average default values for each occupancy type used in Hazard. The default values are the basis for both financial and economic damage calculation in *Flood Damage Wizard*. Structural dollar-per-square-foot values and average square footage values in Table 3 originate in the RSMeans (2006) calculations. Content and inventory unit values in Table 3 find their genesis in the Hazus Earthquake module (FEMA, 1999). *Flood Damage Wizard* uses values in Table 3 to fill gaps in the cadastral data.

Depreciation of each structure was calculated using the specified age of the structure in the cadastre dataset and the depreciation functions from the Hazus R package. These depreciation functions estimate the percentage of value lost for a structure of a given age (HAZUS, 2012a) and are based on the Mean Square Foot Cost Methodology outlined in R.S. Means (2006).

The first step in calculating the financial and economic damages to structures during flooding is calculating the full replacement cost:

$$SFR_i = SQ_i * MSQ_i \quad (4)$$

Here SFR_i is the i th structures full structural replacement value, SQ_i is the i th structures square footage, MSQ_i is the dollar-per-square-foot replacement cost estimate from the fourth column in Table 3. SQ_i can either be supplied by the cadastre dataset or approximated using the third column in Table 3.

Financial damage to the inundated structure of each Address Point-Cadastre site is calculated as

$$DASFR_i = SFR_i * PDS_{i,j} \quad (5)$$

Where $DASFR_i$ is equal to the structural damage at full replacement cost (\$) for the i th building and $PDS_{i,j}$ is the i th structures percentage damage at j th depth. The structural depth-damage relationship is that which *FuzzyWuzzy* determines.

Financial damage to the inundated contents in each Address Point-Cadastre site is calculated as

$$DACFR_i = SFR_i * PDC_{i,j} * PVSV_i \quad (6)$$

Where $DACFR_i$ is equal to the content damage at full replacement cost (\$) for the i th building and $PDC_{i,j}$ is the i th buildings percentage damages at the j th depth. This content depth-damage

relationship is that which *FuzzyWuzzy* determines. $PVSV_i$ equals the percent of structural value that is attributed to content value, which corresponds to fifth column in Table 3.

Financial damage to the inundated inventory in each Address Point-Cadastre site is calculated as

$$DAIFR_i = SQ_i * ASSQ_i * PDI_{i,j} * PVAS_i \quad (7)$$

Where $DAIFR_i$ is equal to the inventory damage at full replacement cost (\$) for the i th building, SQ_i is the i th structures square footage, $ASSQ_i$ is the dollar-per-square-foot annual sales estimate in the sixth column of Table 3, and PDI_i is the i th structures percentage damages at j th depth.

This content depth-damage relationship is that which *FuzzyWuzzy* determines. $PVAS_i$ equals the percent of inventory value that is maintained based upon estimated annual sales, which corresponds to seventh column in Table 3. Inventory damage is assumed to be at full replacement value in both financial and economic terms.

Economic damage to the inundated structure of each Address Point-Cadastre site is calculated as

$$DESFR_i = SFR_i - SFR_i * DE_{i,k} \quad (8)$$

Where $DESFR_i$ is equal to the structural damage at depreciated full replacement cost (\$) for the i th building, SFR_i is the i th structures full structural replacement value, $DE_{i,k}$ is the percentage of depreciation for the i th structure of k th age. If the cadastre datasets does not provide structural age, national averages are used (National Center for Education Statistics, 1999; O'Connor, 2004; Kacker, 2009; U.S. Census Bureau, 2015; U.S. Energy Information Administration, 2015). If the data used provides a market value estimate, $DESFR_i$ is equal to this market value. The authors assume the structural age characteristics of the agriculture occupancy group are consistent with the industrial occupancy group. The authors assume that the age characteristics of the REL1, or

church, occupancy group are consistent with the commercial occupancy groups. Because of the uncertainty associated with this average age estimate, error bands are calculated. These error bands represent the extreme cases were the structure is new or is fully depreciated. Fully depreciated structures are assumed to have an age of 100 years. Each of the depreciation curves used have a maximum age of 100 years. Thus, with or without structural age, a structure reaches full depreciation at 100 years of age.

Economic damage to the inundated contents of each Address Point-Cadastre site is calculated as:

$$DECFR_i = DESFR_i * PDC_{i,j} * PVSV_i \quad (9)$$

Where $DECFR_i$ is equal to the content damage at depreciated full replacement cost (\$) for the i th building, and $PDC_{i,j}$ is the i th buildings percentage damages at the j th depth. This content depth-damage relationship is that which *FuzzyWuzzy* determines. $PVSV_i$ equals the percent of structural value that is attributed to content value, which corresponds to fifth column in Table 3. Equations 4, 5, 6, 7, 8, and 9 are used in four varying scenarios where a probabilistic distribution of total damage is estimated each Address Point-Cadastre location based on 1) structural, content, and inventory damage 2) structural and content damage 3) structural and inventory damage 4) content and inventory damage 5) structural damage 6) content damage 7) inventory damage.

The authors assume that the cadastre data contains no information on first floor height and thus the default heights presumed in the depth-damage relationship are preserved. As indicated in the Further Research section of this manuscript, the model will eventually be linked to an indirect economic consequence assessment model and will eventually estimate both full replacement cost

and depreciated value for economic consequence assessment, as recommended by Merz et al. (2010).

This study will utilize three select county datasets to validate the systems functionality. These counties include Autauga County, Alabama, Tuscaloosa County, Alabama, and Travis County, Texas. Figure 16 depicts these locations.

Table 3. Structural Damage Estimation Parameters. These parameters function to fill in data gaps when no local estimates are available for estimating structural value, content value, and inventory value. From Chapter 14 of the Hazus Technical Manual, excluding column 8. Occupancy Classes RES1 and RES3 have been generalized from the more expanded Hazus versions.

| Occupancy | Description | Means Typical Size (Square Foot) | Means Structure Replacement Value / Square Foot | Content Value (% of Structure Replacement Value) | Annual Sales / Square Foot | Business Inventory (% of Gross Annual Sales) | Average Age |
|--------------------|--|----------------------------------|---|--|----------------------------|--|-------------|
| | | Square Foot | \$/Square Foot | % | \$/Square Foot | % | Years |
| RES1 (Generalized) | Single Family Dwelling | 1,800 | 94.49 | 50 | | | 38 |
| RES2 | Manufactured Housing | 1,625 | 95.75 | 50 | | | 38 |
| RES3 (Generalized) | Multi Family Dwelling | 40,000 | 135.39 | 50 | | | 38 |
| RES4 | Temporary Lodging | 135,000 | 132.52 | 50 | | | 38 |
| RES5 | Institutional Dormitory | 25,000 | 150.96 | 50 | | | 38 |
| RES6 | Nursing Home | 25,000 | 126.95 | 50 | | | 38 |
| COM1 | Retail Trade | 110,000 | 82.63 | 100 | 46 | 13 | 32 |
| COM2 | Wholesale Trade | 30,000 | 75.95 | 100 | 66 | 10 | 32 |
| COM3 | Personal and Repair Services | 10,000 | 102.34 | 100 | | | 32 |
| COM4 | Professional/Technical/Business Services | 80,000 | 133.43 | 100 | | | 32 |
| COM5 | Banks | 4,100 | 191.53 | 100 | | | 32 |
| COM6 | Hospital | 55,000 | 224.29 | 150 | | | 32 |
| COM7 | Medical Office/Clinic | 7,000 | 164.18 | 150 | | | 32 |
| COM8 | Entertainment and Recreation | 5,000 | 170.51 | 100 | | | 32 |
| COM9 | Theaters | 12,000 | 122.05 | 100 | | | 32 |
| COM10 | Parking | 145,000 | 43.72 | 50 | | | 32 |
| IND1 | Heavy Industry | 30,000 | 88.28 | 150 | 616 | 5 | 23 |
| IND2 | Light Industry | 30,000 | 75.95 | 150 | 196 | 4 | 23 |
| IND3 | Food/Drugs/Chemicals | 45,000 | 145.07 | 150 | 602 | 5 | 23 |
| IND4 | Metals/Minerals Processing | 45,000 | 145.07 | 150 | 567 | 3 | 23 |
| IND5 | High Technology | 45,000 | 145.07 | 150 | 378 | 4 | 23 |
| IND6 | Construction | 30,000 | 75.95 | 100 | 664 | 2 | 23 |
| AGR1 | Agriculture | 30,000 | 75.95 | 100 | 128 | 8 | 23 |
| REL1 | Church | 17,000 | 138.57 | 100 | | | 32 |
| GOV1 | General Services | 11,000 | 107.28 | 100 | | | 31 |
| GOV2 | Emergency Response | 11,000 | 166.59 | 150 | | | 31 |
| EDU1 | Schools/Libraries | 130,000 | 115.31 | 100 | | | 42 |
| EDU2 | Colleges/Universities | 50,000 | 144.73 | 150 | | | 42 |

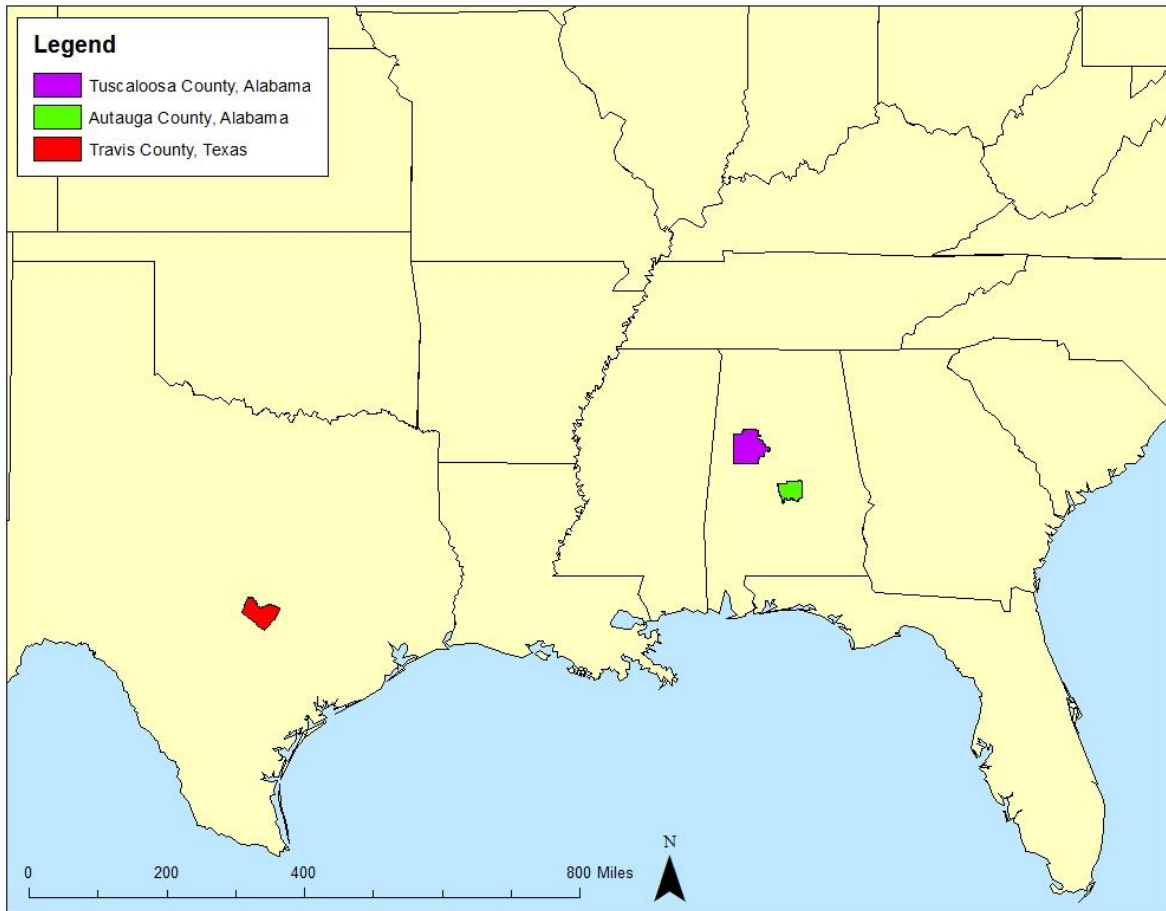


Figure 16. Locations of each Case Study used in this Research

Results

U.S. Army Corps of Engineers Hydrologic Engineering Center River Analysis System (HEC-RAS) models developed by the Alabama Department of Economic and Community Affairs (ADECA) simulated inundation in each of the study areas. Table 4 describes the data provided by each county. Figure 17 depicts a map of each flooded area. Figure 18 represents an example of recorded damage in Autauga County. Autauga County exemplified a standardized used code employed by all Alabama counties (Alabama Department of Revenue, 2005). Additionally, a portion of Autauga County Address Points offer details concerning land use and also describe the condition of the structure at each location for the southeast section of the

county. An equivalent year-depreciation relationship, as described in Table 5 makes use of the condition description. Using the yearly equivalent, *Flood Damage Wizard* approximates a percentage depreciation using the Hazus depreciation curves. For all study locations, missing damage estimations occur in large part because of either unidentified usage descriptions for the address or lack of a structure at the location. The Analysis section of this research further examines the causes of missing damage estimates at Address Point locations.

Table 7 illustrates probabilistic estimates of economic and financial damage in Tuscaloosa County. Table 8 and Table 9 are an example of damages calculated as an average of the probabilistic damage predictions. Normalized Curve and Site Descriptions are those that have had all stop words, extraneous symbols, and capitalized words removed from them. Where no data on square footage, structural age, number of stories, or basement presence are within either the Address Point or cadastral data, *Flood Damage Wizard* uses national averages. Thus, the results from Tuscaloosa County, use only Means Structure Replacement Value/Square Foot from Table 3. Autauga County and Travis County use additional national averages from Table 3.

In most cases, the curve and site matching process is suitable. However, there are exceptions. The site description provided by the Address Point-Cadastre data in Autauga County, in some instances, results in a combined description “single family residential commercial”. This misleading description confuses the *Flood Damage Wizard* text matching algorithm. This results in a damage estimate that is both a commercial/industrial and residential. However, the match value generated by such points lowers because of this error in the dataset. In Tuscaloosa County, two points appear to provide questionable approximate matches. First, the normalized curve descriptions “lock shop”, “auto service”, and “boat service” match with the

point description “service shop low partition one story level”. Secondly, the “post office” curve pairs with the “office two story level” Address-Point Cadastre description. This process also occurs with several Address Point – cadastre descriptions in Travis County, Texas. However, these extensions are not out of the question and to assume that the characteristics of the selected curves are in line with those of the Address Point-Cadastre data is not out of the question. A more detailed site description would provide a better match in this example.

A proportion of the Address Point-Cadastre locations in all three test locations do not carry any description with them. This lack of a description originates from several factors including a lack of coincidental geometry. Figure 19 illustrates a lack of coincidental geometry where the Address Points fall outside the boundaries of the parcel polygon and do not inherit the polygons attributes. However, in the majority of cases, blank record indicates a vacant lot and thus, no structural damage.

Using the inundation grids from each case study, the researchers performed a Level I Hazus Version 2.2 analysis using Level I homogeneous and dasymetric datasets. The researchers compared direct economic building damage that these Hazus analysis generated to the total damage estimates from *Flood Damage Wizard*. Table 10 displays this comparison. The dasymetric datasets are homogeneous Level I datasets with the undeveloped areas removed from them. Thus, FEMA considers the dasymetric datasets a better damage approximator than the homogeneous datasets (FEMA, 2015). Level II datasets were not available in all of the counties utilized in this study, thus the researchers use Level I datasets. Here, *Flood Damage Wizard* is in most cases able to observe the trend of the values that Hazus generates. In general, the overestimation of building damage by Hazus Level I is similar to that which was observed by Ding et al. (2008) when comparing Level I and Level II datasets. The only deviation from this

trend is in Autauga County where the dasymetric Level I and the *Flood Damage Wizard* building damage outputs are nearly equal. This equality may indicate that either the dasymetric data used in this region generates a much better damage estimate than its homogeneous counterpart or that the lack of valuation data in Autauga County’s dataset overstates damage.

Table 4. Pertinent data provided by each county in the study. Note that Autauga County Address Points provide a condition description for part of the county.

| Data | Autauga County, Alabama | Tuscaloosa County, Alabama | Travis County, Texas |
|---------------------------------|---|-----------------------------------|-----------------------------|
| Use Description | X | X | X |
| Square Footage | | X | X |
| Basement Presence | | X | X |
| Number of Stories | | X | X |
| Year Built or Equivalent | Partial condition description included in Address Point dataset | X | |

Table 5. Translation of qualitative structural condition description to yearly depreciation equivalent. These numbers are arbitrary and require additional validation.

| Description | Yearly Depreciation Equivalent |
|---------------------|---------------------------------------|
| Sound | 20 |
| Deteriorated | 75 |

Table 6. Select subset of resultant matches using Flood Damage Wizard in Tuscaloosa County, Alabama.

| Site ID | Occupancy | Class | Depth-Damage Normalized Curve Description | Address Point-Cadastre Normalized Description | Match Value | Water Depth (meters) | Percent Damage |
|---------|-----------|-------|---|---|-------------|----------------------|----------------|
| 1270 | RES1 | Cont | two story w 1 2 living area single family residential dwelling | single family two story level | 88 | 0.61 | 10 |
| 1270 | RES1 | Bldg | two story w 1 2 living area single family residential dwelling | single family two story level | 88 | 0.61 | 4 |
| 29317 | RES2 | Cont | mobile home | mobile home one story level | 100 | 0.30 | 27 |
| 29317 | RES2 | Bldg | mobile home | mobile home one story level | 100 | 0.30 | 19 |
| 29318 | RES1 | Cont | one 1 2 story w 1 2 living area single family residential dwelling | single family one story level | 88 | 0.30 | 6 |
| 29318 | RES1 | Bldg | one 1 2 story w 1 2 living area single family residential dwelling | single family one story level | 88 | 0.30 | 3 |
| 29321 | RES1 | Cont | one 1 2 story w 1 2 living area single family residential dwelling | single family one story level | 88 | 0.30 | 6 |
| 29321 | RES1 | Bldg | one 1 2 story w 1 2 living area single family residential dwelling | single family one story level | 88 | 0.30 | 3 |
| 62748 | IND1 | Inv | lock shop | service shop low partition one story level | 62 | 0.30 | 33 |
| 62748 | COM3 | Cont | auto service | service shop low partition one story level | 74 | 0.30 | 40 |
| 62748 | COM3 | Bldg | boat service | service shop low partition one story level | 74 | 0.30 | 12 |
| 62749 | IND1 | Inv | lock shop | service shop low partition one story level | 62 | 0.30 | 33 |
| 62749 | COM3 | Cont | auto service | service shop low partition one story level | 74 | 0.30 | 40 |
| 62749 | COM3 | Bldg | boat service | service shop low partition one story level | 74 | 0.30 | 12 |
| 63726 | GOV1 | Cont | post office | office two story level | 71 | 1.83 | 100 |
| 63726 | GOV1 | Bldg | post office | office two story level | 71 | 1.83 | 27 |

Table 7. Select probabilistic economic damage estimation from Tuscaloosa County, Alabama.

| Site ID | Structure Percent Damage | Content Percent Damage | Inventory Percentage Damage | Structural Damage | Content Damage | Inventory Damage | Total Damage |
|---------|--------------------------|------------------------|-----------------------------|-------------------|----------------|------------------|--------------|
| Units | % | % | % | \$ | \$ | \$ | \$ |
| 1270 | 4 | 10 | 0 | 12,925 | 16,156 | 0 | 29,079 |
| 1270 | 4 | 10 | 0 | 12,925 | 15,183 | 0 | 28,106 |
| 1270 | 4 | 10 | 0 | 12,925 | 12,652 | 0 | 25,576 |
| 1270 | 4 | 10 | 0 | 12,146 | 16,156 | 0 | 28,301 |
| 1270 | 4 | 10 | 0 | 12,146 | 15,183 | 0 | 27,328 |
| 1270 | 4 | 10 | 0 | 12,146 | 12,652 | 0 | 24,798 |
| 1270 | 4 | 10 | 0 | 10,122 | 16,156 | 0 | 26,276 |
| 1270 | 4 | 10 | 0 | 10,122 | 15,183 | 0 | 25,303 |
| 1270 | 4 | 10 | 0 | 10,122 | 12,652 | 0 | 22,773 |
| 1270 | 4 | 0 | 0 | 12,925 | 0 | 0 | 12,924 |
| 1270 | 4 | 0 | 0 | 12,146 | 0 | 0 | 12,146 |
| 1270 | 4 | 0 | 0 | 10,122 | 0 | 0 | 10,121 |
| 1270 | 0 | 10 | 0 | 0 | 16,156 | 0 | 16,155 |
| 1270 | 0 | 10 | 0 | 0 | 15,183 | 0 | 15,182 |
| 1270 | 0 | 10 | 0 | 0 | 12,652 | 0 | 12,652 |

Table 8. Select estimated financial damages from the Tuscaloosa County, Alabama Address Point and cadastre data.

| Site ID | Depth | Average Percent Building Damage | Average Percent Content Damage | Average Percent Inventory Damage | Average Building Damage | Average Content Damage | Average Inventory Damage | Average Total Damage |
|---------|--------|---------------------------------|--------------------------------|----------------------------------|-------------------------|------------------------|--------------------------|----------------------|
| Units | Meters | % | % | % | \$ | \$ | \$ | \$ |
| 1270 | 0.61 | 3 | 7 | 0 | 10,381 | 12,977 | 0 | 23,357 |
| 29317 | 0.30 | 13 | 18 | 0 | 17,465 | 12,409 | 0 | 29,873 |
| 29318 | 0.30 | 2 | 4 | 0 | 1,482 | 1,482 | 0 | 2,963 |
| 29321 | 0.30 | 2 | 4 | 0 | 3,483 | 3,483 | 0 | 6,965 |
| 62748 | 0.30 | 7 | 23 | 19 | 147,370 | 491,232 | 121,968 | 760,569 |
| 62749 | 0.30 | 7 | 23 | 19 | 147,370 | 491,232 | 121,968 | 760,569 |
| 63726 | 1.83 | 18 | 67 | 0 | 21,628 | 80,102 | 0 | 101,729 |

Table 9. Select estimated economic damages from the Tuscaloosa County, Alabama Address Point and cadastre Data.

| Site ID | Depth | Average Percent Building Damage | Average Percent Content Damage | Average Percent Inventory Damage | Average Building Damage | Average Content Damage | Average Inventory Damage | Average Total Damage |
|---------|--------|---------------------------------|--------------------------------|----------------------------------|-------------------------|------------------------|--------------------------|----------------------|
| Units | meters | % | % | % | \$ | \$ | \$ | \$ |
| 1270 | 0.61 | 3 | 8 | 0 | 9,385 | 11,731 | 0 | 21,115 |
| 29317 | 0.30 | 13 | 18 | 0 | 11,352 | 8,066 | 0 | 19,417 |
| 29318 | 0.30 | 2 | 5 | 0 | 510 | 510 | 0 | 1,019 |
| 29321 | 0.30 | 2 | 5 | 0 | 3,204 | 3,204 | 0 | 6,408 |
| 62748 | 0.30 | 7 | 23 | 19 | 142,949 | 476,495 | 121,968 | 741,411 |
| 62749 | 0.30 | 7 | 23 | 19 | 142,949 | 476,495 | 121,968 | 741,411 |
| 63726 | 1.83 | 18 | 67 | 0 | 19,897 | 73,694 | 0 | 93,591 |

Table 10. Flood Damage Wizard damage compared to Hazus Level I homogeneous and dasymetric analyses.

| | Homogenous Hazus Level I | Dasymetric Hazus Level I | Flood Damage Wizard |
|----------------------------|--------------------------|--------------------------|---------------------|
| | \$ | \$ | \$ |
| Autauga County, Alabama | 41,431,000 | 12,763,000 | 12,655,656 |
| Tuscaloosa County, Alabama | 21,961,000 | 21,863,000 | 2,941,030 |
| Travis County, Texas | 75,001,442 | 74,529,000 | 33,747,231 |

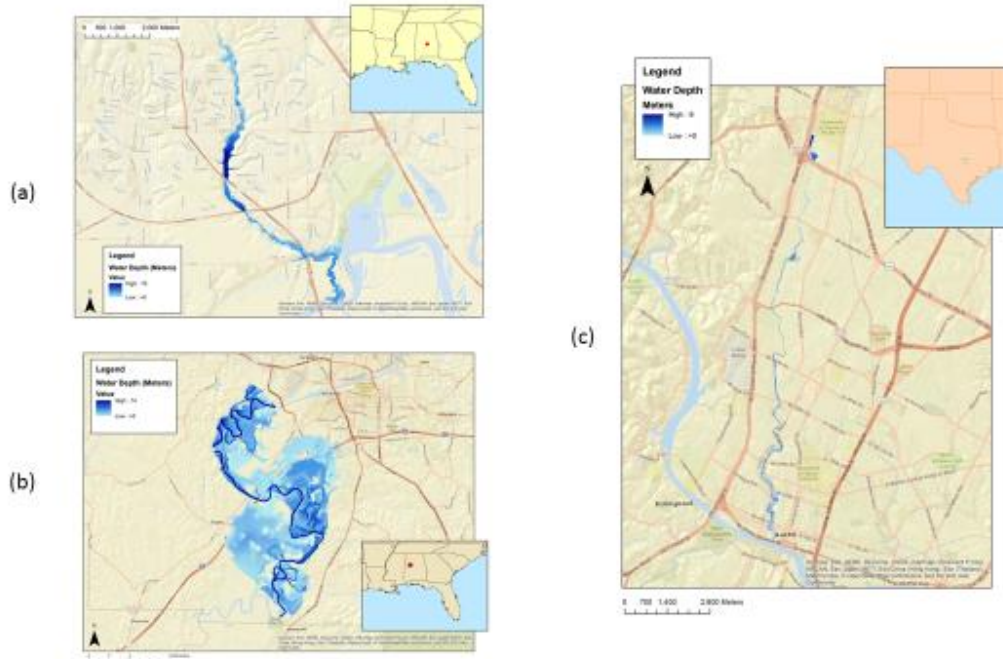


Figure 17. Inundation locations used in this study (a) Pine Creek in Autauga County, Alabama (b) Black Warrior River in Tuscaloosa County, Alabama (c) Shoal Creek in Travis County, Texas.

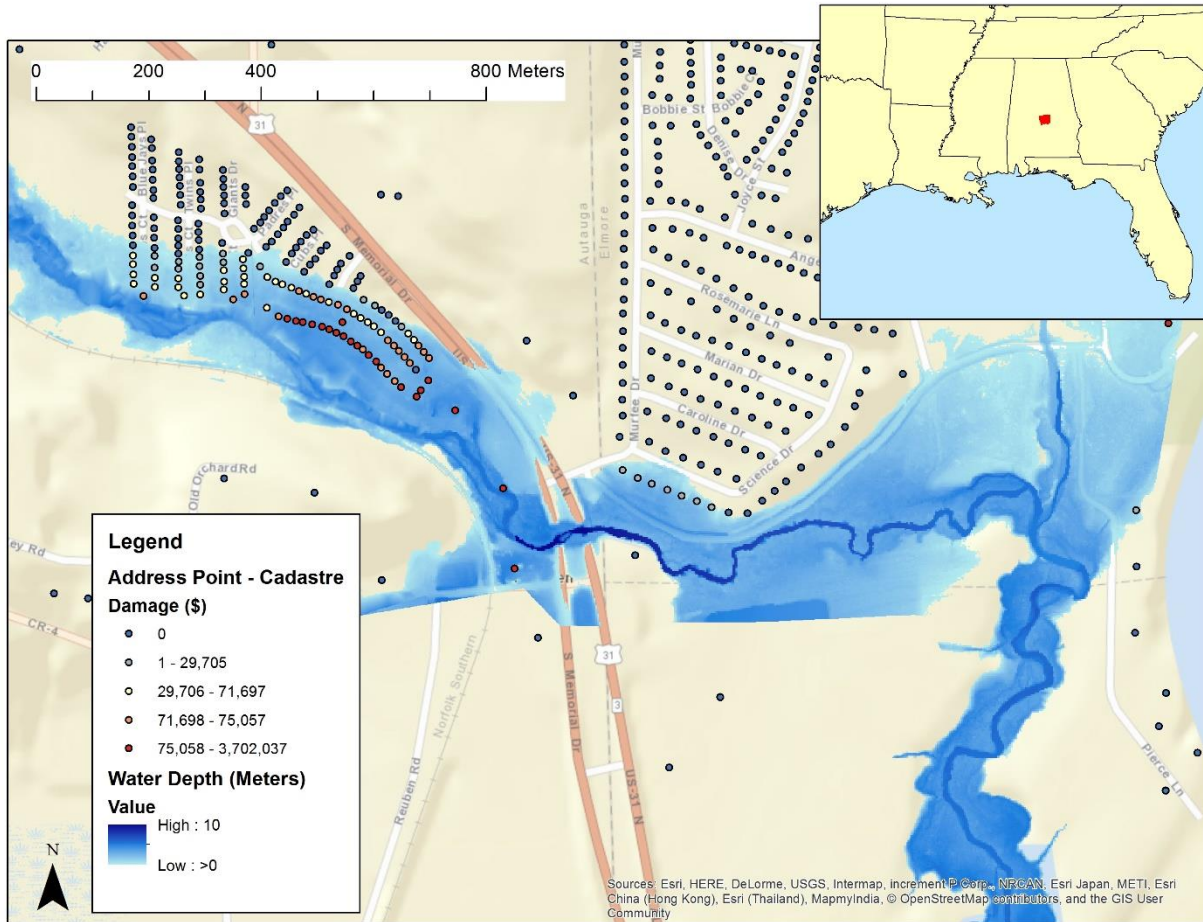


Figure 18. Subset of damage locations at Address Points during 100-Year event along Pine Creek in Autauga County, Alabama.

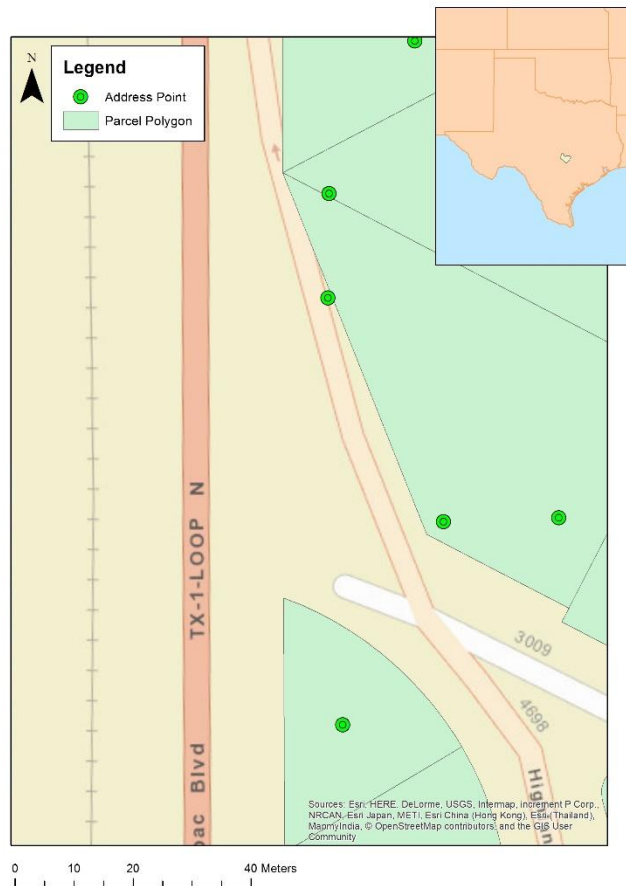


Figure 19. Location where the address points falls outside the parcel polygon in Travis County, Texas. This led to a blank description for the parcel which is unusable by Flood Damage Wizard.

Conclusions

Flood Damage Wizard has proven that it can make use of a variety of Address Point-Cadastre datasets and generate flood damage estimates. The estimates of residential loss appear feasible, considering that the average flood insurance claim, between 2010 and 2014, for residential properties was \$39,000 (NFIP, 2016b). However, the process is iterative and as additional datasets like those in the above research are provided, the more robust *Flood Damage Wizard* will become. As the NFIP (2016a) asserts that the average commercial flood insurance claim was \$89,000, commercial and possibly industrial estimates seem overestimated by *Flood*

Damage Wizard on several occasions. Results indicate that an approximate flood damage study is possible using the *Flood Damage Wizard* toolkit. Supplementing local inputs from the Address Point-Cadastre datasets using data from the Hazus Technical Manual (2012), RSMeans (2006), U.S. Census Bureau (2015), U.S. Energy Information Administration (2015), Kacker (2009), O'Connor (2004), and the National Center for Education Statistics (1999) allows for a flexible and adaptive framework that suits multiple data resolutions. How this compares to other models, such as Hazus Level II analysis, and to real world scenarios will be a part of future study. Further investigation will examine this matter of calibration and validation. However, this initial framework presents an opportunity to use a widely available dataset, which is typically not used in a dynamic nature. This usage would be akin to how the NFIE-Hydro system utilizes High Resolution Rapid Refresh (HRRR) datasets (NOAA, 2016).

Presently, the intention is not for *Flood Damage Wizard* to replace detailed field study. The intent is to produce a system capable of rapid flood damage estimation that can assist decision makers in short lead time scenarios. The length of this lead time finds its basis in the NFIE-Hydro framework. For instance, the user of the system can forecast areas of high damage and preposition resources (i.e., manpower, sand bags, etc.) to impede this damage. If high resolution economic data are available for the geographic area of interest, the researchers suggest that this data be used instead of cadastral and Address Point data.

Even amongst this small sample size the researchers note that an immense amount of heterogeneity exists among different parcel datasets. In the case of Autauga County, Alabama only use descriptions come within the parcel dataset. A subset of the Address Points also provide a use description. Autauga County presents an interesting case in which the Address Point information provides a qualitative description of depreciation. Relating qualitative

depreciation descriptions to quantitative descriptions is a matter of further study. The depreciation figures presented in this paper require further empirical evaluation. The Autauga County data also eludes to the presence of a predefined use code structure for all parcels in Alabama. This development is encouraging, as the minimum input into *Flood Damage Wizard* is a use description. Statewide use descriptions are likely to be present in other states as well. The Travis County, Texas dataset provides more property description than the Autauga County dataset, yet still lacks a year built descriptor. The lack of a year built for the structure resulted in an immense amount of uncertainty where in which the economic value of the structure was particularly imprecise. However, the usage descriptions, square footage, and number of stories for each structure should improve each individual damage estimate. The Tuscaloosa County, Alabama dataset provides description, year built, square footage, basement presence/absence, and number of stories. Thus, the Tuscaloosa County dataset was the most descriptive of those appearing in this research. The resulting damage estimates were more localized than in the case of Autauga County and Travis County.

The results of a *Flood Damage Wizard* run also begin to communicate the uncertainty of flood damage estimation through probabilistic prediction and use of uncertainty bounds in estimated age. *Flood Damage Wizard* assumes two extremes are possible – either the structure is brand new or completely depreciated. Additionally, all residential units have three depreciation functions applied to them. Using this methodology if a user provides only a description of each building location, *Flood Damage Wizard* estimates a probabilistic distribution of at least three financial damage estimates and 15 economic damage estimates. Quantification of the uncertainty resulting from the use of default values will warrant additional investigation. Communication of the uncertainty in these estimates could be provided if both a mean and standard deviation are

found for the default values. These methodologies require further investigation and serve to exemplify how uncertainty can be represented within the model.

Considerable uncertainty exists when using the depth-damage functions. Freni et al. (2010) document this uncertainty in their study and suggest that better estimation of the depth-damage curves will reduce uncertainty. A method of comparison similar to that used by Apel et al. (2009) will provide guidance in how to investigate such uncertainty. Further, the address points constitute approximate building locations. Both the curve utilized and the perceived location of the structure have considerable uncertainty associated with them. Attempts to quantify this uncertainty are an avenue of further research. Address Point locations constitute approximate building locations. Figure 20 demonstrates the approximate nature of Address Points in Autauga County, Alabama. In Figure 20, note that some points fall in line with building locations, while others are quite far away. Utilizing aerial photography in determining building locations may offer a more robust building locator. The work of Ghaffarian and Ghaffarian (2014) may provide a means of better estimating building locations using high-resolution aerial imagery. The use of building footprints in lieu of point locations may also provide better damage estimations. Additional properties of the floodwaters that influence damage, such as flow velocity, may also need to be considered.

Each of the three domains, *FuzzyWuzzy* provides a suitable match between each Address Point-Cadastre site and depth-damage function. Further comparison with other heuristic algorithms, such as expert systems or in data mining approaches such as topic modeling will be explored to optimize the selection process.

In each case study, *Flood Damage Wizard* damage estimates are feasible damage estimations. Further testing is needed to determine how the system compares to Hazus analysis

and/or a real-world scenario. Apel et al. (2009) again provide an example which demonstrates a means by which such a comparison can occur. Further refinement of *Flood Damage Wizard* will be necessary as additional datasets become available.

This research builds upon the precedent Hazus R package (Goteti, 2014) which represents the first example of Hazus in a community development environment. *Flood Damage Wizard* like the Hazus R package is unconstrained in terms of operating system (OS) or third party software dependency. *Flood Damage Wizard* is open source code, implementable in a variety of environments. Thus, as the development of Hazus moves forward, this research intends to serve as a potential model for future iterations of Hazus. Future work will examine how other functions within Hazus can be exhumed. This functionality includes impact on infrastructure such as roads and distribution systems, loss of life, and indirect economic consequence assessment.

The researchers envision this tool primarily as a means of accelerating the response phase of a flood event. A flood damage assessment tool which works in real time could assist in the Preliminary Damage Assessment (PDA). A PDA is undertaken by FEMA at the request of an impacted state's governor to identify the extent of damage during any natural disaster and to determine whether a federal disaster declaration is necessary. At small scales, the PDA process is undertaken by teams of local, state, and federal representatives who survey the impacted area. In large events, this team can be augmented or replaced by technical models (FEMA 2012). The methods presented in this research highlight how technology implementation might expedite the PDA process.

Using predictions of flood damage estimation provide a means by which the emergency planners and responders can communicate risk to the general public. Figure 18 provides an

example of a flood hazard map based on flood damage at the Address Point locations. Further, more advanced analytics can be used within geographic information systems, such as Figure 21. In Figure 21, the Inverse Distance Weighted (IDW) approach generates a raster which serves to further communicate flood risk using forecasted damage estimates (Esri, 2012). These maps are just two of many ways in which the estimation of flood damage can communicate risk in pre-disaster scenarios.

The company KatRisk demonstrates that large scale damage modeling is possible in cluster computing environments. KatRisk utilizes the Titan supercomputer at Oak Ridge National Lab to retrospectively analyze flood damage in a methodology similar to the one envisioned here (KatRisk, 2015). Further analysis of collaboration between the research presented here and the work KatRisk produces will require further exploration.

The cadastral datasets utilized in this study are available in nearly every county in the United States. Though both the Committee on Land Parcel Databases (2007) and Abt Associates Inc. (2013) highlight investigations into a standardized national dataset, none exists in the contemporary. The researchers assert that information within the cadastral datasets is pivotal in understanding the effects of flooding on the built environment and a region's economy. This importance is due to its widespread availability and capacity as a geospatial descriptor of localized building and economic characteristics. As such, this research asserts the value in a national cadastral dataset as part of the Open Water Data Initiative (OWDI). The NFIE was use case 1 for this federal initiative and its intention was to better understand the potential use of OWDI data (Advisory Committee on Water Information, 2015). Thus, this paper concludes that making available a standardized parcel dataset should assist in better understanding flood risk.

Though the fuzzy text matching algorithms employed here provides relevant results, topic models may provide a more robust text matching schema. Topic Models are a data mining approach to interpreting the relationships between strings, usually in the form of structured documents. In topic modelling, a collection of documents or text strings form a collective corpus. Topic modeling assumes that the theme or themes of a document or string closely align to the words that comprise the document or string. Thus, instead of searching for information using a keyword approach, data can be explored using a themed approach. Additionally, multiple themes can and are typically captured in one document (Blei, 2012).

Topic Modeling can assume many forms. The simplest and most common of these topic models is Latent Dirichlet Allocation (LDA). LDA assumes a corpus to have a set of topics and each document within the corpus has a distribution of these topics within its word content. Thus, a portion of a document may fall under topic 1 based on the fact that word 1, word 2, and word 3 are within the document. Topic models infer the hidden structure of a corpus, thus the topics and distribution of words that comprise these topics are initially unknown by the model. The topic model infers the word structure of each topic using the corpus (Blei, 2012).

Topic Models have been applied to mine social media (Lim et al., 2013; Paul and Dredze, 2014) and analyze web supplied user reviews (Rossetti 2015). Essentially, the use of topic models pervades any repository in which information overload is a possibility (Yang et al., 2015). Traditionally, LDA Topic Models categorize structured strings or documents and do not perform well in short, unstructured environments (Zhao et al., 2011; Lim et al., 2013). However, Lim et al. (2013) and Zhao et al. (2011) both prove that they are usable in unstructured environments as well.

Expert systems or other form of artificial intelligence (AI) technology may also provide an alternative means of selecting the appropriate depth-damage relationship for each structure. AI are applicable in a variety of heuristically driven processes. Several decision support system (DSS) applications meld deterministic model outputs with heuristic decision making (Ahmad and Simonovic, 2001; Ahmad and Simonovic, 2006). In this instance, interviews with subject matter experts who select depth-damage curves when constructing Level II Hazus datasets could elicit the rules by which the selection process is undertaken. Many avenues exist in which AI can assist in the selection and modeling processes.

Incorporation of full economic consequence assessment will be pursued using the prior work of Alva-Lizarraga (2013) on economic consequences to drinking water disruptions. The full Hazus methodology will also be explored (FEMA 2012a). Each of the methods utilize the input-output (IO) methodology. The work of Alva-Lizarraga (2013) estimates the impact on an economy which experiences a short-term (30 days or less) loss in water service. An adaptation of the model to flooding scenarios could be achieved by successfully identifying pre- and post-event resiliency actions that a local community can take during a flood event. These resiliency actions would then be parameterized for model inclusion. The work of Keogh et al. (2011) and the questions these researchers asked the residents of Charleville, Australia can be used as a preliminary framework to assess resiliency. Additionally, data compiled by Thieken et al. (2005) in Germany also illustrates resiliency measures taken in the midst of a flooding event. A somewhat different approach is taken in Hazus (FEMA, 2012a), though both utilize IMPLAN data (MIG, 2015) to model indirect economic consequences.

Further evaluation will examine how the model performs in comparison to Hazus-MH and Level II building inventories. Comparison to Hazus Level II datasets can be accomplished

by using the datasets put together for FEMA Flood Risk Reports, such as that developed by the Alabama Department of Economic and Community Affairs (ADECA) Office of Water Resources (OWR) for the Upper Alabama Watershed (ADECA, 2013).

Comparison to actual flooding events and recorded damages will also be pursued. This validation process is very difficult, due to a lack of requisite datasets that describe damage and location (Merz et al., 2010). However, under certain circumstances, models have been successfully examined (Apel et al., 2009). Once a suitable scenario is found, a validation procedure will be undertaken.

The Autauga County, Alabama dataset included a qualitative assessment of each Address Point location. Additional datasets may also have this characteristic. Evaluation of how to use qualitative descriptions and ascertain how they coincide with quantitative depreciation will be another topic of inquiry.

Optimization of the code within *Flood Damage Wizard* is necessary before achieving functionality in a real time environment. Using the local RC2 computing cluster at The University of Alabama, a 12-CPU simulation can take several days with larger flood events in highly developed areas. The cause of this bottleneck is the fuzzy logic, text matching processes occurring within the Python code. Thus, to be useful, *Flood Damage Wizard* require more efficient processing. Optimization can be based on the current fuzzy logic schema or the exploration of more robust tools, such as topic models or expert systems, which will require a learning component. The learning component can be based upon a rule structure (expert systems) or an evolving lookup table. A lookup table which matches common parcel descriptions to the best depth-damage function prior to implementation of the fuzzy logic

algorithm will expedite processing. This lookup table can be incrementally generated using fuzzy matching and the parcel datasets themselves.

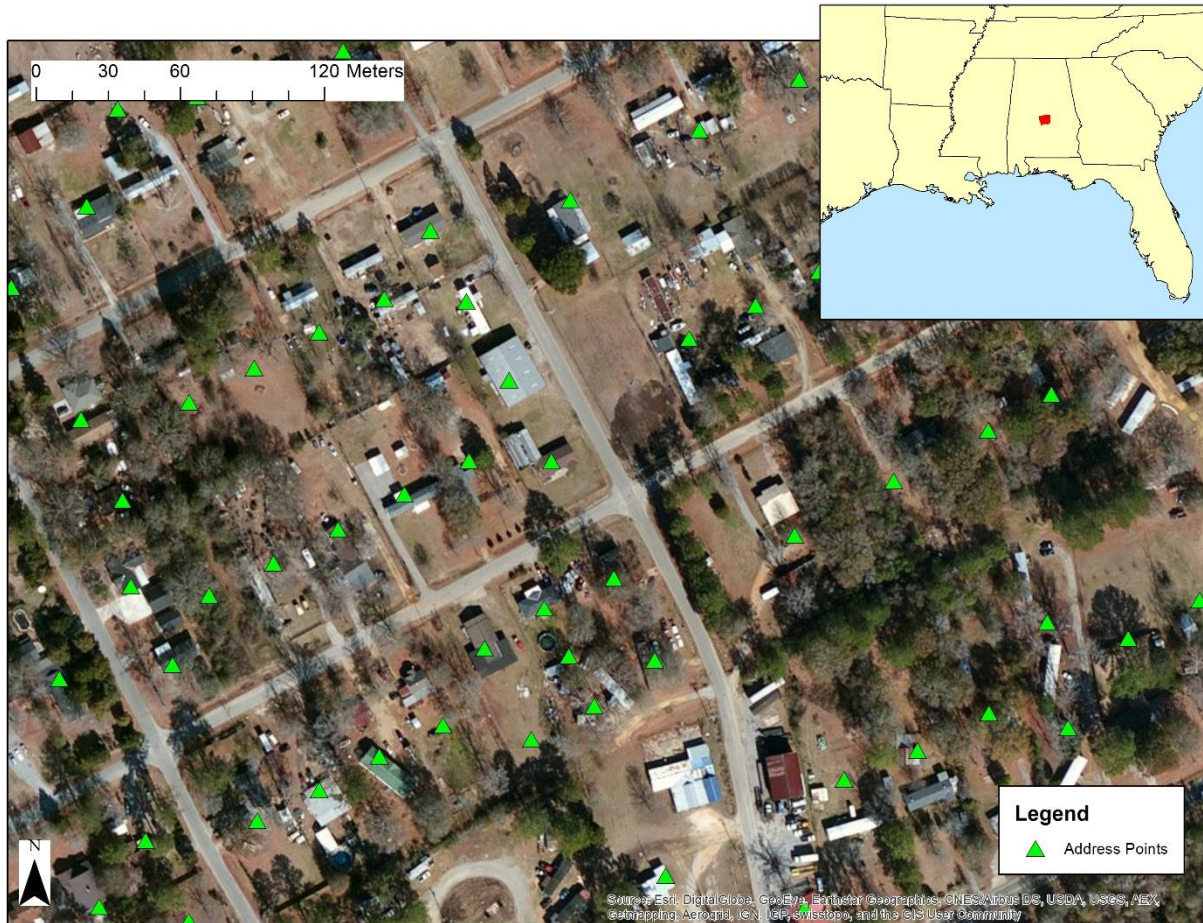


Figure 20. Comparison of Address Point data to aerial imagery. Here it is demonstrated that Address Points are, in most cases, approximate building locations.

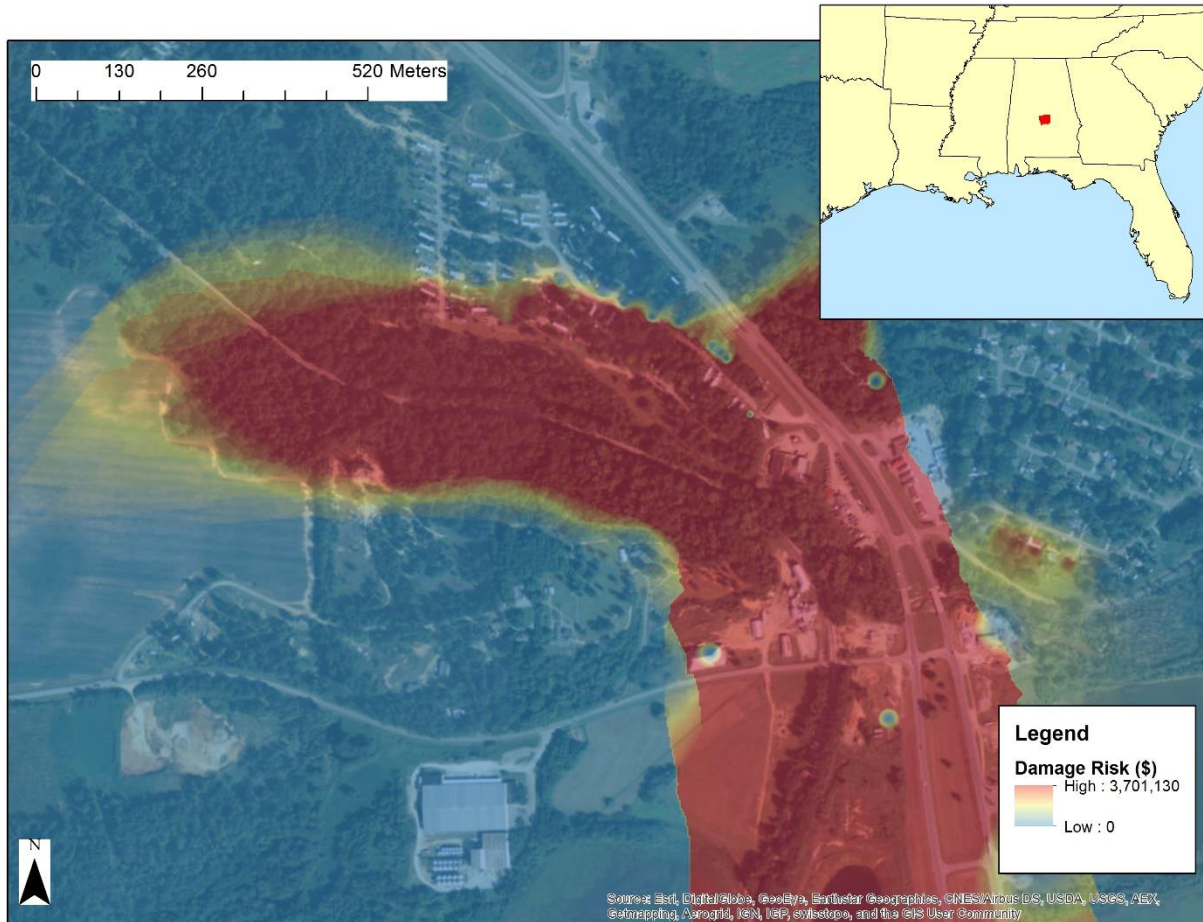


Figure 21. Inverse Distance Weighted (IDW) example of how flood risk can be communicated with forecasted damage estimation.

References

- Abt Associates Inc. (2013). The feasibility of developing a national parcel database: County data records project final report. Bethesda, MD: U.S. Department of Housing and Urban Development.
- Advisory Committee on Water Information. (2015). Open water data initiative overview. Retrieved from <http://acwi.gov/spatial/owdi/>
- Ahmad, S. and S. P. Simonovic. (2001). Integration of heuristic knowledge with analytical tools for the selection of flood damage reduction measures. *Canadian Journal of Civil Engineering*, 28(2), 208-221. DOI: 10.1139/100-097
- Ahmad, S. and S. P. Simonovic. (2006). An intelligent decision support system for management of floods. *Water Resources Management*, 20(3), 391-410. DOI: 10.1007/s11269-006-0326-3

- Alabama Department of Economic and Community Affairs (ADECA). (2013). Flood Risk Report: Upper Alabama Watershed, HUC 03150201.
- Alabama Department of Revenue. (2005). Alabama Appraisal Manual 2005.
- Alva-Lizarraga, S. M. (2013). The Economic Consequences of Water Utility Disruptions. Ph.D. Dissertation, University of Missouri, Columbia, Missouri.
- Apel, H., G. Aronica, H. Kreibich, and A. Thielen. (2009). Flood risk analyses—how detailed do we need to be? *Natural Hazards*, 49(1), 79-98. DOI: 10.1007/s11069-008-9277-8
- Apel, H., A. H. Thielen, B. Merz, and G. Blöschl. (2004). Flood risk assessment and associated uncertainty. *Natural Hazards and Earth System Science*, 4(2), 295-308.
- Bai, Y., H. Zhuang, and D. Wang. (2007). Advanced fuzzy logic technologies in industrial applications. Springer, Natick, Massachusetts, ISBN-13: 978-1-84628-468-7
- Banks, J., J. Camp, and M. Abkowitz. (2014). Adaptation planning for floods: a review of available tools. *Natural Hazards*, 70(2), 1327-1337, doi: 10.1007/s11069-013-0876-7.
- Bennett, R., A. Rajabifard, M. Kalantari, J. Wallace, and I. Williamson. (2010). Cadastral futures: building a new vision for the nature and role of cadastres. In: *XXIV FIG International Congress*. International Federation of Surveyors, pp. 11-16.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84. DOI: 10.1145/2133806.2133826
- Buahin, C., C. Fagan, N. Sangwan, and C. Rae. (2015). Probabilistic Inundation Delineation Using River Cross-Section Rating Curve Library Approach. In: *3rd CUAHSI Conference on HydroInformatics*.
- Chen, H. and N. Chang. (2006). Decision Support for Allocation of Watershed Pollution Load using Grey Fuzzy Multiobjective Programming. *Journal of the American Water Resources Association (JAWRA)* 42(3), 725-745. DOI: 10.1111/j.1752-1688.2006.tb04488.x
- Cohen, A. (2011). FuzzyWuzzy: Fuzzy string matching in python. Retrieved from <http://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-matching-in-python/>
- Committee on Land Parcel Databases. (2007). National land parcel data: A vision for the future. Washington, D.C.: National Academy Press.
- Das, S. and R. Lee. (1988). A Nontraditional Methodology for Flood Stage-Damage Calculations. *Journal of the American Water Resources Association (JAWRA)*, 24(6), 1263-1272.

- Davis, S. A., & Skaggs, L. (1992). Catalog of residential depth-damage functions: Used by the army corps of engineers in flood damage estimation. (No. 92-R-3). Washington, D.C.: U.S. Army Corps of Engineers.
- Ding, A., J. White, P. Ullman, and A. Fashokun. (2008). Evaluation of HAZUS-MH Flood Model with Local Data and Other Program. *Natural Hazards Review*, 9(1), 20-28, doi: 10.1061/(ASCE)1527-6988(2008)9:1(20).
- Eckstein, O. (1958). Water resource development: The economics of project evaluation. Harvard University Press, Cambridge, Massachusetts.
- Esri. (2012). How inverse distance weighted interpolation works. Retrieved from <http://resources.arcgis.com/en/help/main/10.1/index.html#//00310000002m000000>
- Federal Emergency Management Agency (FEMA). (1999). HAZUS: Earthquake Loss Estimation Methodology, Technical Manuals I, II, and III. Washington, D.C.: Federal Emergency Management Agency.
- Federal Emergency Management Agency (FEMA). (2012a). Hazus-MH technical manual. (Computer Program Documentation). Federal Emergency Management Agency (FEMA).
- Federal Emergency Management Agency (FEMA). (2012b). Hazus-MH user manual. (Computer Program Documentation). Washington, D.C.: Federal Emergency Management Agency.
- Federal Emergency Management Agency (FEMA). (2012c). Preliminary damage assessment for individual assistance operations manual. (Standard Operating Procedures (SOP) or Operations Manual No. 9327.2-PR). Washington, D.C.: Federal Emergency Management Agency (FEMA).
- Federal Emergency Management Agency (FEMA). (2015). Hazus online download quick reference guide. Washington, D.C.: Federal Emergency Management Agency (FEMA).
- Flagship GIS, Inc. View our alabama client sites. Retrieved from <http://www.alabamagis.com/>
- Freni, G., G. La Loggia, and V. Notaro. (2010). Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation. *Water Science and Technology*, 61(12), 2979. DOI: 10.2166/wst.2010.177
- Ghaffarian, S. and S. Ghaffarian. (2014). Automatic building detection based on Purposive FastICA (PFICA) algorithm using monocular high resolution Google Earth images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 97, 152-159. DOI: 10.1016/j.isprsjprs.2014.08.017

- Gochis, D. (2015). CONUS-wide evaluation of the operational NFIE prediction system. In: 3rd CUAHSI Conference on HydroInformatics.
- Google. (2015). US parcel data search. Retrieved from <https://support.google.com/earth/answer/2470102?hl=en>
- Goteti, G. (2014). Hazus: Damage functions from FEMA's HAZUS software for use in modeling financial losses from natural disasters. Retrieved from <https://cran.r-project.org/web/packages/hazus/index.html>
- Grigg, N. S. and O. J. Helweg. (1975). State-of-the-Art of Estimating Flood Damage in Urban Areas. *Journal of the American Water Resources Association (JAWRA)*, 11(2), 379-390.
- Huang, Y., & Madey, G. (2004). Web data integration using approximate string join. Paper presented at the Proceedings of the 13th International World Wide Web Conference, New York, New York. 364-365.
- Kacker, P. (2009). Essays on Plant-level Productivity, Market Structure, and Enterprise Growth. Ph.D. Dissertation, American University, Washington, D.C.
- KatRisk. (2015). High resolution catastrophe risk solutions. Retrieved from <http://www.katrisk.com/>
- Kentucky Property Valuation Administrators. (2006). Kentucky assessors. Retrieved from <http://www.kypvas.com/>
- Keogh, D. U., A. Apan, S. Mushtaq, D. King, and M. Thomas. (2011). Resilience, vulnerability and adaptive capacity of an inland rural town prone to flooding: a climate change adaptation case study of Charleville, Queensland, Australia. *Natural Hazards*, 59(2), 699-723. DOI: 10.1007/s11069-011-9791-y
- Lee, G., K. Jun, and E. Chung. (2015). Group decision-making approach for flood vulnerability identification using the fuzzy VIKOR method. *Natural Hazards and Earth System Science*, 15(4), 863-874. DOI: 10.5194/nhess-15-863-2015
- Lim, K. W., C. Chen, and W. Buntine. (2013). Twitter-network topic model: A full Bayesian treatment for social network and text modeling. In: *NIPS2013 Topic Model Workshop*, pp. 4.
- Maidment, D. R. (2014). National Flood Interoperability Experiment. *2014 American Geophysical Union (AGU) Fall Meeting*. American Geophysical Union, San Francisco, California.
- Merz, B., H. Kreibich, R. Schwarze, and A. Thielen. (2010). Review article" Assessment of economic flood damage". *Natural Hazards and Earth System Science*, 10(8), 1697-1724. DOI: 10.5194/nhess-10-1697-2010

- MIG. (2015). Implan. Retrieved from <http://www.implan.com/>
- National Center for Education Statistics. (1999). How old are america's public schools. National Center for Education Statistics.
- National Flood Insurance Program (NFIP). (2013). Technical documentation of NFIP actuarial assumptions and methods. National Flood Insurance Program (NFIP).
- National Flood Insurance Program (NFIP). (2016a). Commercial coverage. Retrieved from https://www.floodsmart.gov/floodsmart/pages/commercial_coverage/cc_overview.jsp
- National Flood Insurance Program (NFIP). (2016b). Residential coverage. Retrieved from https://www.floodsmart.gov/floodsmart/pages/residential_coverage/rc_overview.jsp
- National Oceanic and Atmospheric Administration (NOAA). (2016). The high-resolution rapid refresh (HRRR). Retrieved from <http://ruc.noaa.gov/hrrr/>
- Novák, V., I. Perfilieva, and J. Mockor. (1999). Mathematical principles of fuzzy logic. Springer Science & Business Media, New York, ISBN: 978-1-4613-7377-3
- O'Connor, J. (2004). Survey on actual service lives for North American buildings. In: *Woodframe Housing Durability and Disaster Issues Conference*.
- OnlineSearches. (2015). About assessor and property tax records in texas. Retrieved from <http://publicrecords.onlinesearches.com/Texas-Assessor-and-Property-Tax-Records.htm>
- Oliveri, E. and M. Santoro. (2000). Estimation of urban structural flood damages: the case study of Palermo. *Urban Water*, 2(3), 223-234. DOI: 10.1016/S1462-0758(00)00062-5
- Pappenberger, F., H. L. Cloke, D. J. Parker, F. Wetterhall, D. S. Richardson, and J. Thielen. (2015). The monetary benefit of early flood warnings in Europe. *Environmental Science & Policy*, 51, 278-291. DOI: 10.1016/j.envsci.2015.04.016
- Paul, M. J. and M. Dredze. (2014). Discovering health topics in social media using topic models. *PLoS ONE*, 9(8). DOI: 10.1371/journal.pone.0103408
- Penning-Rowsell, E. C. and T. Wilson. (2003). The benefits of flood and coastal defense: techniques and data for 2003. Flood Hazard Research Center, Middlesex University
- Penning-Rowsell, E., C. Johnson, S. Tunstall, S. Tapsell, J. Morris, J. Chatterton, et al. (2005). The benefits of flood and coastal risk management: a handbook of assessment techniques. Food Hazard Research Center, Middlesex University, ISBN: 1-904750-51-6

- Pita, G., R. Francis, Z. Liu, J. Mitrani-Reiser, S. Guikema, and J. Pinelli. (2011). Statistical Tools for Populating/Predicting Input Data of Risk Analysis Models. In: *Vulnerability, Uncertainty, and Risk*, pp. 468-476.
- Python Software Foundation. (2006). 4.4.2 SequenceMatcher examples. Retrieved from <https://docs.python.org/2.4/lib/sequencematcher-examples.html>
- Python Software Foundation. (2016). Difflib – helpers for computing deltas. Retrieved from <https://docs.python.org/2/library/difflib.html>
- Rees, T. (2014). Taxamatch, an algorithm for near ('Fuzzy') matching of scientific names in taxonomic databases. *PloS One*, 9(9), e107510. doi:10.1371/journal.pone.0107510
- ReportAll. (2015). Parlay parcel data layer. Retrieved from <http://reportallusa.com/products/parlay/>
- Rossetti, M., F. Stella, L. Cao, and M. Zanker. (2015). Analysing User Reviews in Tourism with Topic Models. In: *Information and Communication Technologies in Tourism 2015*, pp. 47-58
- RSMeans. (2006). Square Foot Costs, RS Means Company.
- SeatGeek. (2015). Fuzzywuzzy 0.6.0. Retrieved from <https://pypi.python.org/pypi/fuzzywuzzy/0.6.0>
- Sperling, J. and N. von Meyer. (2013). Collecting Parcel Data from Assessors. *Fair & Equitable*, 3-11.
- Tate, E., C. Muñoz, and J. Suchan. (2015). Uncertainty and Sensitivity Analysis of the HAZUS-MH Flood Model. *Natural Hazards Review*, 16(3), 04014030, doi: 10.1061/(ASCE)NH.1527-6996.0000167.
- Thieken, A. H., M. Müller, H. Kreibich, and B. Merz. (2005). Flood damage and influencing factors: New insights from the August 2002 flood in Germany. *Water Resources Research*, 41(12). DOI: 10.1029/2005WR004177
- U.S. Census Bureau. (2015). 2013 housing profile: United states. (Factsheet).U.S. Census Bureau.
- U.S. Army Corps of Engineers (USACE). (2003). Economic guidance memorandum (EGM) 04-01, generic depth-damage relationships for residential structures with basements. (No. EGM 04-01).U.S. Army Corps of Engineers (USACE).
- U.S. Army Corps of Engineers (USACE). (2008). HEC-FDA flood damage reduction analysis user's manual. (Computer Program Documentation No. CPD-72). Davis, CA: U.S. Army Corps of Engineers (USACE).

- U.S. Army Corps of Engineers (USACE). (2012). HEC-FIA flood impact analysis: User's manual. (Computer Program Documentation No. CPD-81). Davis, CA: U.S. Army Corps of Engineers.
- U.S. Energy Information Administration. (2015). A Look at the U.S. Commercial Building Stock: Results from EIA's 2012 Commercial Buildings Energy Consumption Survey (CBECS). Retrieved from <https://www.eia.gov/consumption/commercial/reports/2012/buildstock/>
- Yang, C. and C. Tsai. (2000). Development of a GIS-Based Flood Information System for Floodplain Modeling and Damage Calculation. *Journal of the American Water Resources Association* (JAWRA), 36(3), 567-577. DOI: 10.1111/j.1752-1688.2000.tb04287.x
- Yang, G., D. Wen, N. Chen, and E. Sutinen. (2015). A novel contextual topic model for multi-document summarization. *Expert Systems with Applications*, 42(3), 1340-1352. DOI: 10.1016/j.eswa.2014.09.015
- Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.
- Zhai, G., T. Fukuzono, and S. Ikeda. (2005). Modeling Flood Damage: Case of Tokai Flood 2000. *Journal of the American Water Resources Association* (JAWRA), 41(1), 77.
- Zhao, W. X., J. Jiang, J. Weng, J. He, E. Lim, H. Yan, et al. (2011). Comparing twitter and traditional media using topic models. In: *33rd European Conference on IR Research*, pp. 338-349.

CHAPTER 4 DEVELOPING REAL-TIME REGIONAL FLOOD DAMAGE DATASETS IN A LARGE SCALE, SHORT LEAD TIME, FLOOD DAMAGE PREDICTION ENVIRONMENT, USING FLOOD DAMAGE WIZARD

Introduction

FEMA's United States Hazard Multi-Hazard (Hazus-MH) is a software commonly used in flood damage analysis. Hazus-MH analyses are scalable, dependent upon available data (Scawthorn et al. 2006). These scales are Level 1, 2, and 3 analyses in the Hazus-MH User's Manual (2012b). Prepackaged within Hazus-MH are national default datasets and workflows that estimate the hydrology, hydraulics, damage losses, and economic losses. An analysis that uses only these built-in components is a Level 1 analysis. Level 2 and 3 analyses augment these datasets to perform more detailed flood damage assessment. For instance, the user can alter the building inventory in the study area using a combination of various datasets, such as parcel level tax assessor information from each county. FEMA strongly encourages that a user update the analyses to be either a Level 2 or Level 3 analysis to improve model performance (FEMA 2012a). The Alabama Department of Economic and Community Affairs (ADECA) developed several Hazus-MH Version 2.1 Level 2 models for use in the Flood Risk Report of the Upper Alabama Watershed (ADECA 2013).

A developed framework, *FDW*, combines Address Point and cadastral datasets for rapid damage estimation (Gutenson et al. 2016). This methodology is the first instantiation of a new mechanism for predicting regional flood damage on a nationwide scale. The researchers define regional to be a resolution comparable to a county-level, with the resulting analysis limited by the predicted hydraulics for streams and rivers or building inventory availability. The Damage assessment methodology follows Chapter 14 of the Hazus-MH Technical Manual (2012a). The framework expands the methodology of Hazus-MH to include probabilistic damage forecasts.

The intention is that the Address Point and cadastral datasets update regularly to ensure an up-to-date damage estimation. The resulting software, *FDW*, is an open source application. *FDW* is portable and scalable to the data and needs of the user. A fuzzy logic algorithm developed by *SeatGeek* (Cohen 2011) determines a suitable depth-damage curve for each inundated property by analyzing each parcels description.

This research considers if rapid, high resolution, large scale flood damage modeling for buildings can approach the results of a detailed regional study. Because of limited information from actual flood events, a detailed regional study such as a FEMA Flood Risk Report (FRR) is the most resolute source of information available to the researchers. This examination will compare two flood damage assessment models, a Hazus-MH Level 2 analysis and *FDW*. *FDW* provides the means of rapidly developing and estimating flood damages. Hazus-MH is a fixture in regional damage modeling studies, such as FRRs. The researchers will consider which inputs are critical in developing an accurate assessment, using the *FDW* approach. The inputs considered are solely building characteristics. The building characteristics considered in this assessment include use description, full replacement value, square footage, number of stories, presence/absence of a basement, and year built. Results provide guidance on what inputs are necessary to approach the accuracy of a detailed Level 2 dataset, usable in FEMA Flood Risk Reports. Hazus-MH Level 2 building inventories from FEMA Flood Risk Reports are typically the most detailed datasets available for flood damage modeling. The application of this methodology occurs with the notion that parcel datasets can exist as a data stream, allowing for utilization of the parcel data in real-time and on-demand. Results indicate that the fuzzy logic approach to utilizing cadastral datasets provides total damage results which are approximate to a Hazus-MH Level 2 datasets when a full replacement value is available. The research also finds

that both a Hazus-MH Level 2 inventory and a cadastral dataset have similar inconsistencies in classifying building occupancy or use. Thus, even the most resolute data sources have associated errors.

Review of Literature

Among the more commonly used and freely accessible flood damage analysis models in the United States are the Hydrologic Engineering Center's Flood Damage Reduction Analysis (HEC-FDA), Flood Impact Assessment (HEC-FIA), and Hazus-MH. Generally, each of these models assess direct damage to residential, commercial, agricultural, and industrial users based upon the depth of floods and associated damage at those depths using depth-damage functions (USACE 2008,2012; FEMA 2012a,b).

Evaluation of these three models occurs in studies conducted by Banks et al. (2014) and Ding et al. (2008). Banks et al. (2014) performed a comprehensive review of available software packages geared towards flood planning and found that Hazus-MH was the only one of the 11 reviewed to meet all of the authors' evaluation criteria. This review included HEC-FIA. Ding et al. (2008) assessed Hazus-MH's performance at both Level 1 and Level 2 performance and compared this against a previously conducted study which utilized HEC-FDA. The results of Ding et al. (2008) indicate that the outputs of the Level 2 Hazus-MH simulation were very similar to those created in HEC-FDA, but required a fraction of the man-power used to develop the latter. One of the most advantageous pieces to the Hazus-MH software is its robust default inventories which greatly reduce data acquisition.

Amongst these three models, Hazus-MH is the only one explicitly capable of modeling indirect economic impacts. Indirect impacts are the ripple effects that occur within an economy after significant disruptions to normality. Examining the inter-sectorial dependencies within an

economy describes these indirect effects. In most cases, including Hazus-MH, indirect impact analysis occurs through input/output (IO) modelling (FEMA 2012a). Hazus-MH is also the only one of the three software packages which carry with it a comprehensive database of predefined infrastructural and residential inventories. These inventories describe categories of structures and how they are affected by floods at specific depths (FEMA 2012a). In fact, HEC-FIA has the built-in ability to directly import Hazus-MH infrastructure inventories (USACE 2012).

Hazus-MH analyses are scalable, depending upon available data. These scales are Level 1, 2, and 3 analyses in the Hazus-MH User's Manual (2012). Level 1 analysis uses only the default inventories and workflows. Specific economic inventories are available for download for each U.S. state. According to the Hazus-MH User's Manual (FEMA 2012b), a generalized Level 1 analysis is not sufficient enough for detailed flood damage analysis. In particular, the User's Manual indicates that the inventory described in a Level 1 Hazus-MH analysis will not be able to account for individual and regional heterogeneity among structures in a given community. Thus, a user should update the analyses to be either a Level 2 or more detailed Level 3 analysis.

In the literature, attempts to create flood damage estimation models focus on developing long range planning tools. Jonkman et al. (2008) developed a robust direct and indirect flood damage economic consequence assessment tool. The methodology Jonkman et al. (2008) developed focused on coupling hydraulics and economic modeling to gridded land use datasets. As in Hazus-MH, HEC-FDA, and HEC-FIA, flood damage assessment occurs through depth-damage relationships for each type of land use. Dutta et al. (2003) employ another grid based approach using stage-damage relationships which coupled hydrology with flood damage estimation.

Other methods examine damage estimation methodologies beyond the use of just depth-damage relationships. Middlelmann-Fernandes (2010) found that using a combination of stage-damage or velocity-stage-damage functions produce more realistic damage estimates. Kreibich et al. (2009) also recommend the use of velocity-damage functions under certain scenarios. Additionally, other variables, such as energy head, can impact actual flood damage (Pistrika 2010). However, water depth is the single greatest predictor of flood damage to buildings (Kreibich et al. 2009). Most of these parameters mentioned depend upon the complexity of the physical model utilized.

Tate et al. (2014) consider both physical and economic model inputs in their sensitivity analysis of flood damage to residential buildings. Tate et al. (2014) find that the accuracy of the digital elevation model (DEM) is the most influential input selection on model outputs in Hazus-MH modeling. However, Tate et al. (2014) report that residential building input parameters, such as the structures value, also have significant influence on model uncertainty. Srestha (2014) examine the importance of building inputs in Hazus-MH modeling and find that first floor heights and foundation types are the most significant influence on building damages. Srestha (2014) notes that square footage, building age, and construction type have little influence on Hazus-MH outputs. De Moel & Aerts (2011) find that the depth-damage curve and value of each property are the greatest sources of uncertainty and that uncertainty in the physical model (i.e., the inundation grid) will have to be large to significantly influence damage estimates.

Hydrology and hydraulics are moving into the realm of continental scale simulation and forecasting. This large-scale application of the water cycle presents an opportunity to expand the platform of flood damage estimation. The National Flood Interoperability Experiment (NFIE) held at the National Water Center (NWC) in Tuscaloosa, Alabama led to the development of this

large scale flow and inundation forecasting framework capable of 15-hour lead-time forecasts for the entire coterminous United States (CONUS). This concept will become the National Water Model.

FDW is a new Geographic Information System (GIS) based system designed by the authors which utilizes Address Point and cadastral information to estimate flood damage to buildings, both financial and economic. *FDW* is written completely in Python 2.7. *FDW* is an open source application, which is portable and scalable to the data and needs of the user. Unlike Hazus-MH, HEC-FIA, or HEC-FDA, the user can script analyses on the computational platform of their choosing. *FDW* can work with many property description datasets and building location descriptions converted into GIS-based point shapefiles of building locations. The only data input requirement from the point shapefile is a use description for each building. The intent of *FDW* is to rapidly construct building inventories and assess damage in short lead-time scenarios (Gutenson et al. 2016).

Many prior studies address flood damage estimation in the context of long term proactive planning. The long-term resiliency projects brought about by the FEMA Hazard Mitigation Grants provide a \$5.10 of benefit for every dollar spent (Rose et al. 2007). The impact of these grants can be analyzed in Hazus-MH. In short time-frame scenarios, however, FEMA Hazard Mitigation Grants are not feasible resiliency options and in many locations high resolution Hazus-MH Level 2 or 3 analyses are not available. The issue of damage estimation and associated benefits in short-term flood scenarios is not a focus in past research. The lack of research in this arena is primarily due to the sparcity and spatial inconsistency of short-term hydrologic forecasts and the assumption that the assemblage of adequate building inventories cannot occur within small time frames.

Ding et al. (2008) estimate that it can take 16 hours to assemble a Hazus-MH Level 2 dataset. In a short time frame, the time costs associated with producing and utilizing the data are prohibitive. Thus, the datasets used in detailed studies will not likely be available due to both time and cost considerations. The Carsell et al. (2004) system is capable of estimating flood damage with hydrometeorological forecasts but a prior Comprehensive Study provides an understanding of the economics of the region. This Comprehensive Study required “rigorous data collection”. Nastev et al. (2015) detail a set of GIS-based tools from Canada which facilitate a rapid building inventory development. However, the Nastev et al. (2015) tool relies upon manual configuration of building parameters through visual survey. Thus, in many areas of the country, if no detailed economic analyses have been undertaken or if no expert is able to development the building inventory, no such damage estimation is feasible.

FDW offers a means to rapidly estimate flood damages for use in expediting the determination of federal disasters which may be useful in improving short-term flood resiliency. Emergency responders may be able to employ resources to prevent substantial financial losses if the locations of these losses are known prior to the floods onset. In addition, following a significant flood event, Preliminary Damage Assessments (PDAs) assess the validity of declaring a federal disaster. A PDA is undertaken by FEMA at the request of an impacted state’s governor, to identify the extent of damage during any natural disaster and to determine the whether a federal disaster declaration is necessary. On small scales, the PDA process is undertaken by teams of local, state, and federal representatives who survey the impacted area. In large events, where on-foot survey is prohibitive, this team can utilize technical models (FEMA 2012c). An important component of this evaluation is the estimate of per capita damage to structures and resulting economic damage. Prior to such events, *FDW* may provide a means to

expedite the PDA development process by estimating damage or providing officials guidance on heavily impacted areas.

Method

This research seeks to determine what inputs from parcel or cadastral datasets are critical for rapidly estimating flood damage using *FDW*. The researchers investigate this by comparing model outputs to those estimated by the Hazus-MH Version 2.1 Flood Module Level 2 building inventories in three counties in the Upper Alabama watershed. The methodology is chosen due to the difficulty in attaining measured flood damage estimates from real world events. Thus, Hazus-MH and the Level 2 building inventory utilized represent the most resolute data available for comparison. Because most of the estimation process employed by *FDW* seeks to estimate building replacement value, the researchers consider two scenarios: 1) where building replacement value is estimated 2) where building replacement value is available. We consider the value parameter of the Level 2 inventories to be the actual market value.

Damage Estimation

Table 11. Default values used by Flood Damage Wizard when the user dataset does not specify them.

| Occupancy | Description | Means Typical Size (Square Foot) | Means Structure Replacement Value / Square Foot | Content Value (% of Structure Replacement Value) | Annual Sales / Square Foot | Business Inventory (% of Gross Annual Sales) | Average Age |
|-----------------------|---|---|---|---|----------------------------------|---|----------------|
| | | Square Foot | \$/Square Foot | % | \$/Square Foot | % | Years |
| RES1 (Generalized) | Single Family Dwelling | 1,800 | 94.49 | 50 | | | 38 |
| RES2 | Manufactured Housing | 1,625 | 95.75 | 50 | | | 38 |
| RES3 (Generalized) | Multi Family Dwelling | 40,000 | 135.39 | 50 | | | 38 |
| RES4 | Temporary Lodging | 135,000 | 132.52 | 50 | | | 38 |
| RES5 | Institutional Dormitory | 25,000 | 150.96 | 50 | | | 38 |
| RES6 | Nursing Home | 25,000 | 126.95 | 50 | | | 38 |
| COM1 | Retail Trade | 110,000 | 82.63 | 100 | 46 | 13 | 32 |
| COM2 | Wholesale Trade | 30,000 | 75.95 | 100 | 66 | 10 | 32 |
| COM3 | Personal and Repair Services | 10,000 | 102.34 | 100 | | | 32 |
| COM4 | Professional/Technical/Business Services | 80,000 | 133.43 | 100 | | | 32 |
| COM5 | Banks | 4,100 | 191.53 | 100 | | | 32 |
| COM6 | Hospital | 55,000 | 224.29 | 150 | | | 32 |
| COM7 | Medical Office/Clinic | 7,000 | 164.18 | 150 | | | 32 |
| COM8 | Entertainment and Recreation | 5,000 | 170.51 | 100 | | | 32 |
| COM9 | Theaters | 12,000 | 122.05 | 100 | | | 32 |
| COM10 | Parking | 145,000 | 43.72 | 50 | | | 32 |
| IND1 | Heavy Industry | 30,000 | 88.28 | 150 | 616 | 5 | 23 |
| IND2 | Light Industry | 30,000 | 75.95 | 150 | 196 | 4 | 23 |

| | | | | | | | |
|------|----------------------------|---------|--------|-----|-----|---|-----------------|
| IND3 | Food/Drugs/Chemicals | 45,000 | 145.07 | 150 | 602 | 5 | 23 |
| IND4 | Metals/Minerals Processing | 45,000 | 145.07 | 150 | 567 | 3 | 23 |
| IND5 | High Technology | 45,000 | 145.07 | 150 | 378 | 4 | 23 |
| IND6 | Construction | 30,000 | 75.95 | 100 | 664 | 2 | 23 |
| AGR1 | Agriculture | 30,000 | 75.95 | 100 | 128 | 8 | 23 ¹ |
| REL1 | Church | 17,000 | 138.57 | 100 | | | 32 ² |
| GOV1 | General Services | 11,000 | 107.28 | 100 | | | 31 |
| GOV2 | Emergency Response | 11,000 | 166.59 | 150 | | | 31 |
| EDU1 | Schools/Libraries | 130,000 | 115.31 | 100 | | | 42 |
| EDU2 | Colleges/Universities | 50,000 | 144.73 | 150 | | | 42 |

1

¹ The assumption is that the characteristics of the agriculture occupancy group are consistent with the industrial occupancy group.

² The assumption is that the characteristics of this occupancy group are consistent with the commercial occupancy groups per U.S. Energy Information Administration (2015).

Method of Comparison

A Level II Hazus analysis exists for each of the five counties which lie adjacent to the Alabama River. Included in this Level II analyses are local building inventories for each of the counties. Three of the five counties have allowed access to their cadastral information. Thus, with the development of *Flood Damage Wizard*, a comparison between a Level II structural inventory and the inventory constructed using *Flood Damage Wizard* is possible with the three acquired datasets. The authors utilize the spatial locations of buildings within the Hazus Level II datasets to ensure 1:1 comparison.

AMEC developed Hydrologic Engineering Center's River Analysis System (HEC-RAS) inundation grids for particular return periods along the Upper Alabama River (AMEC, "Alabama-Coosa River Flood Study" unpublished report for Alabama Department of Economic and Community Affairs (ADECA)). In this research, the authors compare the estimated damage from the Hazus Level II building inventory dataset and *Flood Damage Wizard* and cadastral damage estimates using these inundation grids. The authors use the 10-, 25-, 50-, 100-, and 500-year return period events along the Upper Alabama River in three counties with Hazus Level II data. These counties are Autauga County, Lowndes County, and Montgomery County. This process generates 15 samples for comparison. The researchers consider both composite and individual total damage in these scenarios.

Study Region

Table 12 summarizes the resolution of the data from each county in the study region and the quick facts from the U.S. Census Bureau concerning each counties economic composition. These data present a heterogeneous mix of resolution and composition. The most resolute data source is the Montgomery County data providing all data utilized by *Flood Damage Wizard*. On

the opposite end of the spectrum, Autauga County provides only use description. However, the researchers note that a number of parameters come from both the larger and more populous counties (Montgomery County) and the smaller and socioeconomically challenged locations (Lowndes County).

The Hazus Level II dataset built by ADECA in their 2013 study provides the hydraulic and economic estimates of damage for 10-, 25-, 50-, 100-, and 500-year events. These datasets, in turn, produce an estimate of damage in percentage and dollar amounts for the structure, content, and inventory at each user defined facility in each county. This study compares total dollar damage estimates from each Hazus Level II model to those which are recorded at the same locations by *Flood Damage Wizard*. As a market value estimation, the researchers use the cost value available within the Hazus Level II datasets.

The researchers consider two scenarios. The first is one where no estimate of market value is available but all other inputs provided are available. The inputs' square footage and the year built approximate market value with an estimate of depreciated full replacement value. The second scenario is where a combination of both market value and all model inputs are available. These dual scenarios exist to determine the influence estimating depreciated market value has on accuracy in the modeling process.

Table 12. Data Resolution for Each Parcel Dataset Used in this Study (U.S. Census Bureau, 2015).

| County | Autauga County, Alabama | Lowndes County, Alabama | Montgomery County, Alabama |
|--|-------------------------|-------------------------|----------------------------|
| 2014 Population Estimate (U.S. Census Bureau 2015) | 55,395 | 10,580 | 226,189 |
| Per Capita Income 2013 Dollars (U.S. Census Bureau 2015) | \$24,571 | \$18,368 | \$24,975 |
| Hazus Value Estimate from Level II User Defined Facilities | X | X | X |
| Use Description | X | X | X |
| Number of Stories | | X | X |
| Square Footage | | | X |
| Year Built | | | X |
| Basement Presence | | X | X |

Results

In each of the following scenarios, total damage comparisons omit points where *Flood Damage Wizard* did not generate a damage estimate to compare against the Hazus damage estimates. The reasons for this lack of damage estimates is found in the Discussion section of this research.

Scenario 1: Estimation of Market Value

In this scenario, the researchers used all inputs available from Table 12, except market value. Table 13 presents the total damages estimated by Hazus and the Mean and Max *Flood Damage*

Wizard methods. In this scenario, results generated by *Flood Damage Wizard* are quite different when compared to each Hazus Level II simulation, except in the case of Lowndes County. However, upon examining the point records for each county, as in Figure 22, little correlation exists between *Flood Damage Wizard* and Hazus total damage estimates for each building. In Montgomery County, even when all model inputs are available, the damages estimated are different than those estimated using the Hazus and the Level II dataset.

Table 13. Comparison of total building damage estimates using all values except market value. Note that in cases where area with development, such as Autauga and Montgomery County, estimates in relation to Hazus are much greater.

| | Autauga | | | Lowndes | | | Montgomery | | |
|----------|--------------|---------------|---------------|--------------|--------------|--------------|---------------|---------------|---------------|
| | Hazus | FDW Mean | FDW Max | Hazus | FDW Mean | FDW Max | Hazus | FDW Mean | FDW Max |
| 10-Year | \$2,849,524 | \$42,175,398 | \$97,027,387 | \$1,218,515 | \$1,314,892 | \$2,623,863 | \$32,639,172 | \$12,999,180 | \$21,373,509 |
| 25-Year | \$5,931,777 | \$75,714,581 | \$173,986,304 | \$2,967,019 | \$2,708,046 | \$5,436,875 | \$80,152,354 | \$24,821,475 | \$41,979,215 |
| 50-Year | \$8,396,339 | \$105,987,796 | \$243,610,400 | \$4,886,386 | \$5,486,709 | \$11,234,091 | \$129,097,702 | \$36,829,122 | \$61,995,080 |
| 100-Year | \$12,303,398 | \$143,582,514 | \$328,900,781 | \$6,823,275 | \$8,113,826 | \$16,688,251 | \$231,020,704 | \$77,726,180 | \$129,551,439 |
| 500-Year | \$33,639,080 | \$366,893,480 | \$838,719,108 | \$11,054,138 | \$17,571,175 | \$36,912,297 | \$330,880,118 | \$101,284,922 | \$171,329,623 |

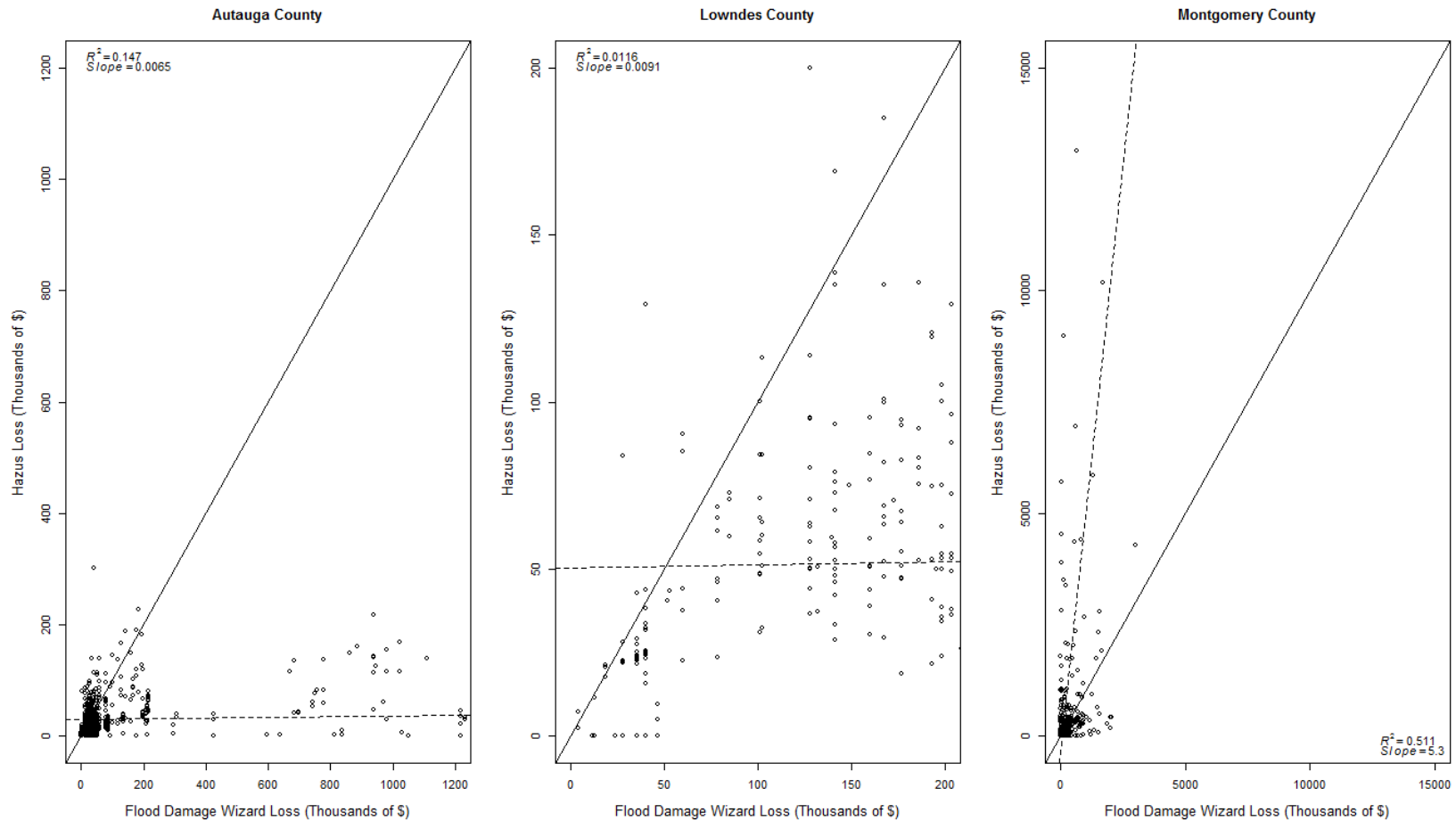


Figure 22. Point total damage estimate comparison between Flood Damage Wizard maximum values and Hazus Level II analysis for each counties 500-year event. Here, the authors use no market value and there seems to be no correlation between Flood Damage Wizard and Hazus total losses. The solid line represents a slope of one and the dashed line is the regression line. When the structural age is unknown, as in Autauga and Montgomery Counties, the maximum value is the full replacement value and Flood Damage Wizard overestimates loss. When structural age is known, as in Montgomery County, Flood Damage Wizard underestimates value and resulting flood loss.

Scenario 2: Market Value Provided

With the addition of market value, the researchers remove considerable uncertainty involved with estimating depreciated full replacement value and calculating the resulting damages. In comparison to Table 13, the researchers note the improvement in *Flood Damage Wizard* total damage estimates, which converge to the values predicted by Hazus in all three counties in Table 14. With the inclusion of market value, both the mean and max *Flood Damage Wizard* total damage estimates are more approximate to the Hazus results than those in Table 13. Table 14 demonstrates that the mean and max *Flood Damage Wizard* methods provide approximate lower and upper bound estimates of damage. In the case of Autauga County, the mean method is more approximate to the Hazus estimate. In the case of Lowndes and Montgomery Counties, the max method is more approximate to Hazus. Figure 23 demonstrates a point comparison of total loss in Autauga, Lowndes, and Montgomery County for each counties 500-year event. Figure 23 demonstrates substantial improvement in correlation. However, Autauga County still performs poorly. This reflects in the results from Table 14, where *Flood Damage Wizard* total damage estimates are less conforming to the Hazus Level II results. The researchers present an argument for the cause of this mismatch at the individual level in the Discussion section of the document.

Table 15 and Table 16 list a general comparison of the match between the curve description and parcel description for single family residences in Autauga and Montgomery Counties, respectively. In Table 15 and Table 16, *Flood Damage Wizard* does not use an inventory curve at these locations. This match is suitable in both instances. Similarly, Table 17 demonstrates additional matches that are suitable. However, other matches appear to be less suitable (Table 18). For instance, the use of the curve “Small Grocery” for the parcel description “Small Tract” in Autauga County is not an appropriate match. Personal correspondence with the Autauga

County Tax Assessor reveals that this class of property has a rather ambiguous definition of use where the property can be residential, commercial, or industrial. A discussion of this error occurs in the Discussion section of this research.

Table 14. Total damage comparison between Hazus and each of the Flood Damage Wizard (FDW) calculations. Note that this comparison omits Hazus locations where FDW was unable to find a suitable depth-damage function.

| | Autauga | | | Lowndes | | | Montgomery | | |
|----------|--------------|--------------|--------------|--------------|-------------|--------------|---------------|---------------|---------------|
| | Hazus | FDW Mean | FDW Max | Hazus | FDW Mean | FDW Max | Hazus | FDW Mean | FDW Max |
| 10-Year | \$2,849,524 | \$4,143,551 | \$6,723,643 | \$1,218,515 | \$696,571 | \$1,018,153 | \$32,641,077 | \$22,166,294 | \$34,503,296 |
| 25-Year | \$5,931,777 | \$7,789,596 | \$12,676,857 | \$2,967,019 | \$1,531,311 | \$2,247,491 | \$80,158,100 | \$44,971,083 | \$73,112,578 |
| 50-Year | \$8,396,339 | \$11,600,740 | \$18,969,754 | \$4,886,386 | \$2,697,073 | \$3,961,115 | \$129,105,891 | \$68,160,750 | \$111,299,468 |
| 100-Year | \$12,303,398 | \$17,538,775 | \$28,799,460 | \$6,823,275 | \$3,784,946 | \$5,562,868 | \$231,029,718 | \$156,971,345 | \$258,047,014 |
| 500-Year | \$44,567,144 | \$50,710,355 | \$84,441,444 | \$11,054,138 | \$7,179,918 | \$10,594,029 | \$332,610,799 | \$210,616,007 | \$339,526,176 |

Table 15. Example of residential single family match in Autauga County dataset. Note that no inventory class was given to this match.

| Building ID | Occupancy | Class | Depth-Damage Curve Description | Parcel Description |
|-------------|-----------|----------|--|--------------------|
| US000001 | RES1 | Content | one & 1/2 story w/ 1/2 living area below | single family |
| US000001 | RES1 | Building | one & 1/2 story w/ 1/2 living area below | single family |

Table 16. Example of residential single family match in Montgomery County dataset. Note that no inventory class was given to this match.

| Building ID | Occupancy | Class | Depth-Damage Curve Description | Parcel Description |
|-------------|-----------|----------|------------------------------------|---|
| 1 | RES1 | Content | two story w/ 1/2 living area below | single family residential two story level |
| 1 | RES1 | Building | two story w/ 1/2 living area below | single family residential two story level |

Table 17. Example of depth-damage curve and description matches made by Flood Damage Wizard in Montgomery County.

| Building ID | Occupancy | Class | Depth-Damage Curve Description | Parcel Description |
|-------------|-----------|-----------|---|----------------------|
| 26 | COM2 | Inventory | Warehouse | Warehouse |
| 26 | COM2 | Contents | Warehouse | Warehouse |
| 26 | COM2 | Building | Warehouse | Warehouse |
| 30 | COM8 | Contents | Fast Food Restaurant, contents, fresh water, short duration | Restaurant Fast Food |
| 30 | COM8 | Building | Fast Food Restaurant, contents, fresh water, short duration | Restaurant Fast Food |

Table 18. Example of questionable depth-damage curve and description matches made by Flood Damage Wizard in Autauga County.

| Building ID | Occupancy | Class | Depth-Damage Curve Description | Parcel Description |
|-------------|-----------|-----------|--------------------------------|-------------------------------|
| US000081 | COM1 | Inventory | Small Grocery | Small Tract |
| US000081 | COM1 | Contents | Small Grocery | Small Tract |
| US000081 | COM1 | Building | Small Grocery | Small Tract |
| US000441 | IND1 | Inventory | Lock Shop | Service Shop Low Partition |
| US000441 | COM3 | Contents | Auto Service | Service Shop Low Partition |
| US000441 | COM3 | Building | Boat Service | Service Shop Low Partition |

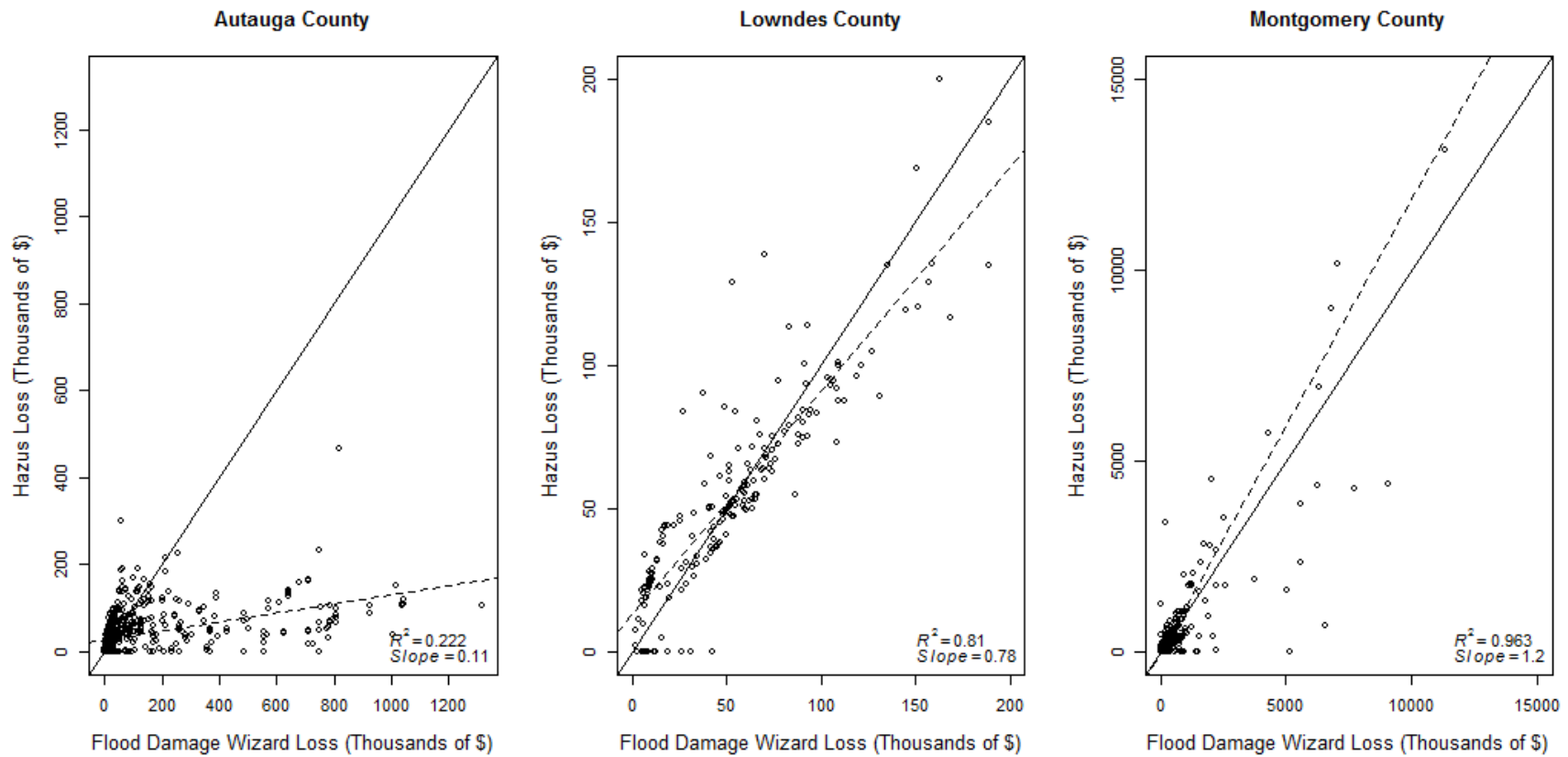


Figure 23. Point total damage estimate comparison between Flood Damage Wizard maximum estimate and Hazus Level II analysis for each counties 500-year event. Here, the use of market value occurs and two of three instances seem to provide similar results. The solid line represents a slope of one and the dashed line is the regression line.

Conclusions

Flood Damage Wizard can match Hazus Level II total damage estimates and individual point damage estimates based on Figure 23. *Flood Damage Wizard* can produce accurate estimates of total damage when an estimate of market value or its equivalent is available. Large differences at the individual point level, such as those in Autauga County, occur primarily because of mismatches between the occupancy listed in the Hazus Level II building inventory and the parcel information utilized.

The simplest explanation for data discrepancies is differences in age between the cadastral and Hazus datasets. In this regard, the Level II Hazus analysis is static and provides only a single glimpse at the geographic layout of the communities. With the use of maintained cadastral datasets, the building inventories constructed are likely a better approximation of reality. The fact that these data are web-delivered in most cases asserts their use as a service, much like the National Water Information System (NWIS) serves out U.S. Geological Survey (USGS) gage information as part of the Open Water Data Initiative (Goodall et al., 2008; Blodgett et al., 2015; OWDI).

Table 19 provides a more thorough comparison of the mismatch between what the Hazus Level II data presents and what *Flood Damage Wizard* asserts to be the use of each building. In this table, the authors visually inspect these locations within Google Maps to approximate the use of each. In this table, the research considers a subset of Hazus Level II buildings classified as RES1 (single family residential) in which *Flood Damage Wizard* predicts that an inventory damage occurs at this location. The prediction of inventory damage insinuates that an improper match occurs as residential buildings do not contain an inventory. Table 19, however, presents that *Flood Damage Wizard* may not be incorrect under all circumstances. In some instances,

Hazus Level II is correct; sometimes *Flood Damage Wizard* is correct, and other times neither are correct. Table 20 also considers a subset of Hazus classified COM1 or retail trade buildings where again this pattern repeats. Figure 24 illustrates the Google Street View® of Point 131 which the Hazus Level II building inventory classifies as a single family residential location; the parcel data classifies as a service shop, and the Google Street View® depicts as a church. Thus, both the Hazus Level II dataset and the cadastral data have associated uncertainty in their occupancy classification.

This comparison alludes to the difficulty in identifying one use for a building. Buildings can have several uses and narrowing this use to one is difficult. The differentiation in assigned usage affects damage modeling primarily in classifying whether a building contains an inventory or not. Variability of inventory value amongst different industrial and commercial uses alone is quite substantial as column 5 and 6 in Table 11 detail. Further, assigning an occupancy of industrial or commercial to a residential location can cause even more severe damage estimation errors. Because, square footage is important to estimating inventory values in both Hazus and *Flood Damage Wizard*, the authors suggest that its use is necessary to approximate a detailed flood damage study. This study finds that there is a great deal of uncertainty associated with the use of either a Hazus Level II building inventory or the cadastral information.

At this moment, the intention is not for *Flood Damage Wizard* to replace detailed field study. The intent is to produce a system capable of rapid flood damage estimation that can assist decision makers in short lead time scenarios. The length of this lead time being determined by the outputs of the National Water Model. For instance, the user of the system can forecast areas of high damage and initiate the FEMA PDA process to determine if federal disaster declaration is necessary. If high resolution economic data are available for the geographic area of interest,

the researchers suggest the use of detailed building inventories instead of cadastral and/or Address Point data.

Here, the authors demonstrate that *Flood Damage Wizard* total damage estimates can approach the total damage results of a Hazus Level II analysis both as composite and point estimates. The data requirements of the approximate approach are a cadastral dataset containing an estimate of market value, use description, and square footage. The selection of depth-damage function is approximate to those selected in a normal Level II Hazus analysis. More complete datasets which include first floor height are likely to yield results closer to the Level II Hazus analysis.

The authors omit cases where matches between the Hazus and cadastral data were not found. One cause for these omissions is incomplete data provided by the county. For example, the data provided by Lowndes County is incomplete and additional data are available, but were not provided by county representatives. In other cases, a sufficient match between the description provided by the parcel data and the depth-damage function description was not found. In other instances, the Hazus Level II point data constituting the building inventory did not share coincidental geometry with the cadastral data. The authors will consider solutions for these omissions.

The mismatches between the Hazus and cadastral data can lead to large differences in individual building locations due to mis-assignment of inventories to buildings. For example, assigning an IND1 inventory curve to a 2,000 square church would lead to an overestimate of \$61,600, if the inventory receives 100% damage. These types of mismatches occur primarily in Autauga and Montgomery Counties where a much larger proportion of the building inventory is industrial or commercial. The damage estimation methodology used here assumes that not all

industrial and commercial sectors possess an inventory. Thus, a minor mismatch can have large estimation consequences.

Considering the results in Figure 23, the researchers caution utilizing solely a cadastral dataset. Indeed, a substantial difference between estimates produced by *Flood Damage Wizard*, and Hazus exists here based on vague cadastral descriptions and other factors previously discussed. The *Flood Damage Wizard* modeling framework is dependent upon sufficient description of buildings for a region. If vague descriptions, such as “Small Tract” are present, modeling with this approach will lead to substantial errors. A means of thoroughly communicating the uncertainty in this classification is necessary.

A secondary requirement of correct building usage classification is its usage in economic impact study that follows damage estimation. In order to understand regional impacts of flood damage and disruption, the proportions of each industry must be understood. For example, the work of Alva-Lizarraga (2013) identifies an example of the input/output methodology used in economic impact analysis. In her work, Alva-Lizarraga uses a temporal estimation of the proportion of each industry impacted by a drinking water disruption in service. The proportion of each economic sector impacted by a flood event is necessary for estimating the direct and indirect economic impact that the disaster supplies.

The most important input from this analysis is that damage estimates are dependent upon accurate market value or economic value estimates. This contribution can come within the parcel datasets themselves, estimated from the values provided by the parcel datasets, or some other proprietary source, such as Zillow or Trulia. Again, all these sources exist as web-services and may function as a service, if the owner offers proper access.

Table 19. Subset of buildings classified as RES1 in Hazus Level II database in Montgomery County, Alabama. In this instance it seems that in both Flood Damage Wizard and the Hazus data mis-classify the data in some instances. Similar circumstances occur in the Autauga County dataset.

| ID | Flood Damage Wizard Occupancy | Parcel Description | Use Determined from Google Maps |
|-----|-------------------------------|----------------------------|---------------------------------|
| 6 | IND1/COM3 | Service shop low partition | Single family residential |
| 131 | IND1/COM3 | Service shop low partition | Church |
| 656 | IND1/COM3 | Service shop low partition | Service Shop |

Table 20. Subset of buildings classified as COM1 in Hazus Level II database in Autauga County, Alabama. In this instance it seems that in both Flood Damage Wizard and the Hazus data mis-classify the data in some instances.

| ID | Flood Damage Wizard Occupancy | Parcel Description | Use Determined from Google Maps |
|------|-------------------------------|--------------------|---------------------------------|
| 582 | COM1 | Retail store | Retail Store |
| 138 | COM1 | Retail mixed | Mobile Home |
| 1088 | RES1 | Single family | Single Family |

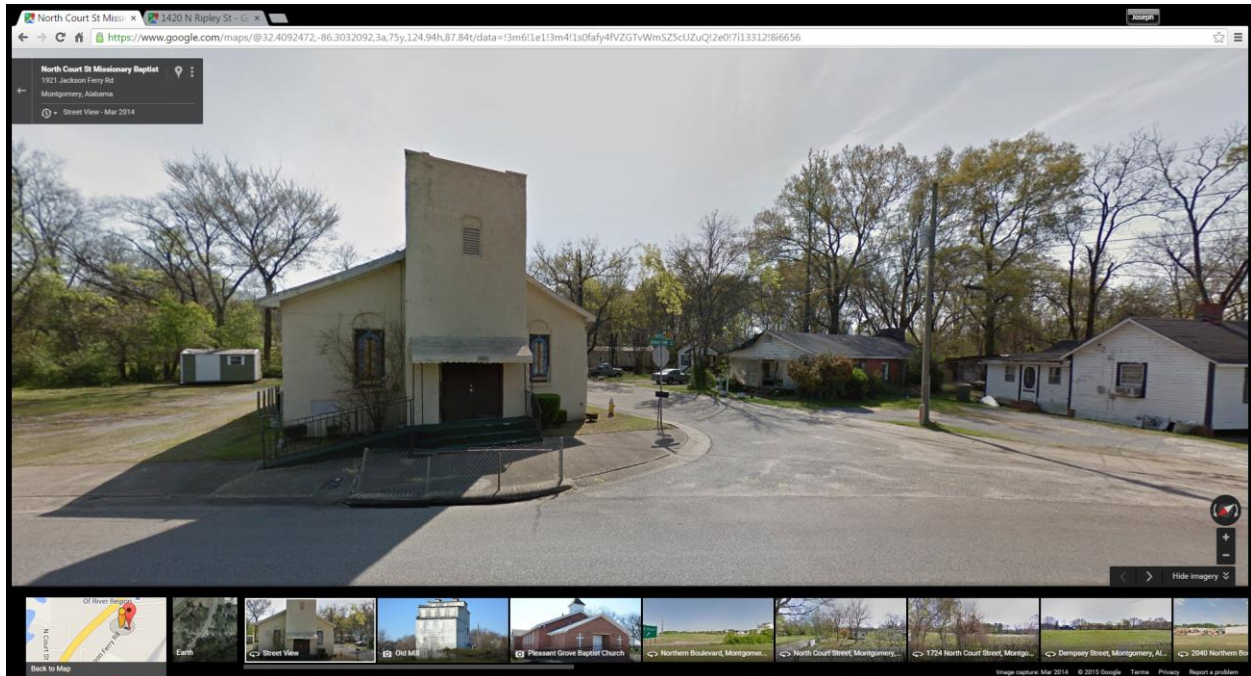


Figure 24. Google Street View® of Point 131 in Montgomery County. The parcel data classifies this as a service shop, the Hazus Level II building inventory classifies the site as single family residential, and from Google Street View®, the site appears to be a church

Future Work

The Literature discusses alternative methods of building inventory assemblage. Wieland et al. (2012) consider the use of omnidirectional image analysis in determining building characteristics. Another approach might consider the use of existing Building Information Models (BIM) in damage analysis. BIM models may become more of an option in creating a building inventory as their use proliferates.

To avoid mismatches between parcel descriptions and depth-damage descriptions, it may benefit the researchers to consider a weighted input scheme. For example, consider the “Small Grocery” curve description paired with “Small Tract” parcel description in Table 18. The word “Small” has very little influence on the proper match in either the parcel or curve description. The inputs “Grocery” and “Tract” are much more significant in determining a proper match. The

current text matching schema matches by assigning equal weights to all terms in a particular text string. A method by which to weight inputs were “Small” is weighted less than either “Grocery” or “Tract” would likely produce a much better text match. *Flood Damage Wizard* assigns a match value to each text string match best fit and will disregard any matches that do not meet a minimum match value (Gutenson et al., 2016). Such a technique could lower the match value score assigned in cases such as this. A lower match value score would then be at least ignored by *Flood Damage Wizard*.

Considerable uncertainty associated with flood damage and loss estimations is present in any methodology employed. Freni et al. (2010) conclude that considerable uncertainty resides with the depth-damage functions. Freni et al. assert that better estimation of these relationships should produce more certain estimations. The work of de Moel and Aerts (2011) supports the work of Freni et al. (2010) and finds that depth-damage relationships account for an additional factor of two uncertainty in their study. Downton and Pielke (2005) note that in their research they observe considerable inaccuracy in damage estimates at local scales. This inaccuracy is somewhat resolved at higher levels of spatial aggregation. Future work will look to measure these uncertainties and produce a system that ultimately reduces them. Additionally, like the work of Merz et al. (2013), future research should consider the use of alternative methodologies in evaluating flood damage beyond depth-damage relationships.

An interesting possibility is the combination of this modeling capacity with Voluntary Geographic Information (VGI). VGI provides a platform for crisis mapping (Middleton et al. 2014). Poser and Dransch (2010) discuss the use of VGI in updating the physical inundation of a flood event and how it can improve flood damage estimates. Du et al. (2015) pursued the use

of VGI in understanding flood warning response and how to improve probabilistic inundation maps.

This methodology only applies to estimation of building damage. The estimation of damage to infrastructure is under pursuit to provide the ability to better understand flood impacts. For example, the use of transportation infrastructure will analyze physical damage to the infrastructure and its effect on the local populace. The use of VGI also applies in this scenario (Schnebele et al. 2014).

The national parameters in Table 11 may not be functionally suitable to approach the resolution of a Hazus Level II analysis. However, more localized default parameters available in each state might approach more realistic outcomes. Resources such as each states Department of Revenue or Building Commission may provide more localized evaluations of these parameters. A developed building inventory would also need to include information on the distribution of default parameters, such as standard deviation. The inclusion of standard deviation would allow for damage models to incorporate the uncertainty which follows the use of such data.

Sawada et al. (2014) present the Urban Rapid Assessment Tools (URAT) which takes advantage of the visualization capabilities of the Google Maps API. The use of Google Maps or Google Earth APIs in this damage analysis may provide better results. This information is proprietary and only limited usage is granted without acquiring usage fees. However, the business and industrial information within this data is tremendous. The researchers envision a tool where the inundation grid defines the search radius and each damage curve provides the search terms within the radius. This methodology would take advantage of not only the data but other text matching and processing capabilities that Google provides. Applications within the Google framework have typically only used the services provided to map natural disaster

outcomes (Laituri and Kodrich 2008). However, Alam et al. (2012) developed a building inventory through a visual survey of Google Maps for use in estimating seismic damage. Akbarzadeh and Wilmot (2015) identify service routes available to evacuees during Hurricane events. OpenStreetMap and Bing Maps also provide a similar API, though they have no usage fees (Schelhorn et al. 2014).

References

- Akbarzadeh, M., and Wilmot, C. G. (2014). Time-Dependent Route Choice in Hurricane Evacuation. *Nat. Hazards Rev.*, 16(2), 04014021. DOI: 10.1061/(ASCE)NH.1527-6996.0000159
- Alabama Department of Economic and Community Affairs (ADECA). (2013). Flood risk report: Upper Alabama watershed, HUC 03150201. (No. 001). Washington, D.C.: U.S. Federal Emergency Management Agency (FEMA).
- Alam, M. N., Tesfamariam, S., and Alam, M. S. (2012). GIS-Based Seismic Damage Estimation: Case Study for the City of Kelowna, BC. *Nat. Hazards Rev.*, 14(1), 66-78. DOI: 10.1061/(ASCE)NH.1527-6996.0000082
- Alva-Lizarraga, S. M. (2013). The Economic Consequences of Water Utility Disruptions. Doctor of Philosophy. University of Missouri—Columbia.
- Banks, J., Camp, J., and Abkowitz, M. (2014). Adaptation planning for floods: a review of available tools. *Nat. Hazards*, 70(2), 1327-1337. DOI: 10.1007/s11069-013-0876-7
- Blodgett, D., Lucido, J., and Kreft, J. (2015). Progress on water data integration and distribution: a summary of select US Geological Survey data systems. *J. Hydroinf.*
- Carsell, K. M., Pingel, N. D., and Ford, D. T. (2004). Quantifying the benefit of a flood warning system. *Nat. Hazards Rev.*, 5(3), 131-140. DOI: 10.1061/(ASCE)1527-6988(2004)5:3(131)
- Cohen, A. (2011). FuzzyWuzzy: Fuzzy string matching in python. Retrieved from <http://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-matching-in-python/>
- de Moel, H., and Aerts, J. (2011). Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates. *Nat. Hazards*, 58(1), 407-425. DOI: 10.1007/s11069-010-9675-6

- Ding, A., White, J., Ullman, P., and Fashokun, A. (2008). Evaluation of HAZUS-MH Flood Model with Local Data and Other Program. *Nat. Hazards Rev.*, 9(1), 20-28. DOI: 10.1061/(ASCE)1527-6988(2008)9:1(20)
- Downton, M. W., and Pielke, R. A. (2005). How accurate are disaster loss data? The case of US flood damage. *Nat. Hazards*, 35(2), 211-228. DOI: 10.1007/s11069-004-4808-4
- Dutta, D., Herath, S., and Musiaka, K. (2003). A mathematical model for flood loss estimation. *Journal of Hydrology*, 277(1), 24-49. DOI: 10.1016/S0022-1694(03)00084-2
- Federal Emergency Management Agency (FEMA). (2012a). Hazus-MH technical manual. (Computer Program Documentation). Federal Emergency Management Agency (FEMA).
- Federal Emergency Management Agency (FEMA). (2012b). Hazus-MH user manual. (Computer Program Documentation). Washington, D.C.: Federal Emergency Management Agency.
- Federal Emergency Management Agency (FEMA). (2012c). Preliminary damage assessment for individual assistance operations manual. (Standard Operating Procedures (SOP) or Operations Manual No. 9327.2-PR). Washington, D.C.: Federal Emergency Management Agency (FEMA).
- Freni, G., La Loggia, G., and Notaro, V. (2010). Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation. *Water Science and Technology*, 61(12), 2979. DOI: 10.2166/wst.2010.177
- Goodall, J. L., Horsburgh, J. S., Whiteaker, T. L., Maidment, D. R., and Zaslavsky, I. (2008). A first approach to web services for the National Water Information System. *Environmental Modelling & Software*, 23(4), 404-411. DOI: 10.1016/j.envsoft.2007.01.005
- Gutenson, J. L., Ernest, A. N. S., Oubeidillah, A. A., Zhu., L., and Sadeghi, S. T. (2016). Rapid flood damage assessment using public domain cadastral and address point data with fuzzy logic algorithms. Manuscript submitted for publication.
- Jonkman, S., Bočkarjova, M., Kok, M., and Bernardini, P. (2008). Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. *Ecol. Econ.*, 66(1), 77-90. DOI: 10.1016/j.ecolecon.2007.12.022
- Kreibich, H., Piroth, K., Seifert, I., Maiwald, H., Kunert, U., Schwarz, J., Merz, B., and Thieken, A. (2009). Is flow velocity a significant parameter in flood damage modelling? *Natural Hazards and Earth System Science*, 9(5), 1679-1692. DOI: 10.5194/nhess-9-1679-2009
- Laituri, M., and Kodrich, K. (2008). On line disaster response community: People as sensors of high magnitude disasters using internet GIS. *Sensors*, 8(5), 3037-3055. DOI: 10.3390/s8053037

- Maidment, D. R. (2014). National Flood Interoperability Experiment. *2014 American Geophysical Union (AGU) Fall Meeting*. American Geophysical Union, San Francisco, California.
- Merz, B., Kreibich, H., and Lall, U. (2013). Multi-variate flood damage assessment: a tree-based data-mining approach. *Natural Hazards and Earth System Science*, 13(1), 53-64. DOI: 10.5194/nhess-13-53-2013
- Merz, B., Kreibich, H., Schwarze, R., and Thielen, A. (2010). Review article "Assessment of economic flood damage". *Natural Hazards and Earth System Science*, 10(8), 1697-1724. DOI: 10.5194/nhess-10-1697-2010
- Middelmann-Fernandes, M. (2010). Flood damage estimation beyond stage–damage functions: an Australian example. *Journal of Flood Risk Management*, 3(1), 88-96. DOI: 10.1111/j.1753-318X.2009.01058.x
- Middleton, S. E., Middleton, L., and Modafferi, S. (2014). Real-time crisis mapping of natural disasters using social media. *Intelligent Systems, IEEE*, 29(2), 9-17. DOI: 10.1109/MIS.2013.126
- Nastev, M., Nollet, M., Abo El Ezz, A., Smirnoff, A., Ploeger, S. K., McGrath, H., Sawada, M., Stefanakis, E., and Parent, M. (2015). Methods and Tools for Natural Hazard Risk Analysis in Eastern Canada: Using Knowledge to Understand Vulnerability and Implement Mitigation Measures. *Nat. Hazards Rev.*, B4015002. DOI: 10.1061/(ASCE)NH.1527-6996.0000209
- Oliveri, E., and Santoro, M. (2000). Estimation of urban structural flood damages: the case study of Palermo. *Urban Water*, 2(3), 223-234. DOI: 10.1016/S1462-0758(00)00062-5
- Pappenberger, F., Cloke, H. L., Parker, D. J., Wetterhall, F., Richardson, D. S., and Thielen, J. (2015). The monetary benefit of early flood warnings in Europe. *Environ. Sci. & Policy*, 51 278-291. DOI: 10.1016/j.envsci.2015.04.016
- Pistrika, A. K., and Jonkman, S. N. (2010). Damage to residential buildings due to flooding of New Orleans after hurricane Katrina. *Nat. Hazards*, 54(2), 413-434. DOI: 10.1007/s11069-009-9476-y
- Poser, K., and Dransch, D. (2010). Volunteered geographic information for disaster management with application to rapid flood damage estimation. *Geomatica*, 64(1), 89-98.
- Sawada, M., Ploeger, K., Elsabbagh, A., Nastev, M., Saatcioglu, M., & Rosetti, E. (2014). Integrated desktop/mobile GIS application for building inventory. (OPEN FILE 7345). Geological Survey of Canada.

- Scawthorn, C., Flores, P., Blais, N., Seligson, H., Tate, E., Chang, S., Mifflin, E., Thomas, W., Murphy, J., and Jones, C. (2006). HAZUS-MH flood loss estimation methodology. II. Damage and loss assessment. *Nat. Hazards Rev.*, 7(SPECIAL ISSUE: Multihazards Loss Estimation and HAZUS), 72-81. DOI: 10.1061/(ASCE)1527-6988(2006)7:2(72)
- Schelhorn, S., Herfort, B., Leiner, R., Zipf, A., and de Albuquerque, J. P. (2014). Identifying Elements at Risk from OpenStreetMap: The Case of Flooding. In: *Proceedings of the 11th International ISCRAM Conference*, pp. 508-512.
- Schnebele, E., Cervone, G., and Waters, N. (2014). Road assessment after flood events using non-authoritative data. *Natural Hazards and Earth System Science*, 14(4), 1007-1015. DOI: 10.5194/nhess-14-1007-2014
- Shrestha, S. (2014). Sensitivity of Hazus-MH Flood Loss Estimates to Selection of Building Parameters: Two Illinois Case Studies. Master's Thesis, Southern Illinois University – Carbondale.
- Tate, E., Muñoz, C., and Suchan, J. (2015). Uncertainty and Sensitivity Analysis of the HAZUS-MH Flood Model. *Nat. Hazards Rev.*, 16(3), 04014030. DOI: 10.1061/(ASCE)NH.1527-6996.0000167
- U.S. Army Corps of Engineers (USACE). (2008). HEC-FDA flood damage reduction analysis user's manual. (Computer Program Documentation No. CPD-72). Davis, CA: U.S. Army Corps of Engineers (USACE).
- U.S. Army Corps of Engineers (USACE). (2012). HEC-FIA flood impact analysis: User's manual. (Computer Program Documentation No. CPD-81). Davis, CA: U.S. Army Corps of Engineers.
- U.S. Census Bureau. (2015). State and county quickfacts. Retrieved from <http://quickfacts.census.gov/qfd/index.html>
- Wieland, M., Pittore, M., Parolai, S., Zschau, J., Moldobekov, B., and Begaliev, U. (2012). Estimating building inventory for rapid seismic vulnerability assessment: Towards an integrated approach based on multi-source imaging. *Soil Dyn. Earthquake Eng.*, 36 70-83. DOI: 10.1016/j.soildyn.2012.01.003

CHAPTER 5 CONCLUSIONS

Research Questions and Conclusions

Research Question 1

This research investigates the role of hydraulic geometries in large scale channel estimation for unregulated streams in Alabama. This investigation was the first of its kind for the entire state of Alabama. The results show a positive correlation between depth, width, and cross-sectional area, calculated by hydraulic geometries equations and those derived from HEC-RAS simulations. Considerable differences in prediction accuracy occur between different physiographic regions in the State of Alabama. The East Gulf Coastal Plain and Piedmont Upland regions of Alabama demonstrate positive correlation between the HEC-RAS dimensions and those predicted by hydraulic geometries. The hydraulic geometries assume that the relationships are uniformly asymptotic, which is conceivably representative of the geomorphology of streams in areas where human alteration to the channel dimensions and watershed has been minimal. Estimating stream channels with hydraulic geometries where development pervades the watershed or where the channel has been subjected to anthropogenic alteration is quite uncertain. In such cases, hydraulic geometries are inaccurate representations of the actual channel dimensions, as they are no longer subject to natural geomorphic restrictions.

The estimated depth of both below and above bankfull discharge hydraulic geometries correlates well with the depth outputs of HEC-RAS in the East Gulf Coastal Plain and Piedmont Upland physiographic regions. This finding is substantial, as it is the only instance found in the literature where researchers use a discharge different than bankfull discharge. The sensitivity of depth to changes in discharge from these regions' streams may originate in the hydraulic geometry coefficients for this physiographic province. Hydraulic geometry estimates of depth in the East Gulf Coastal Plain and Piedmont Upland physiographic region also correlate well with HEC-RAS outputs under both high and low flow conditions, with correlation improving for higher flows. This research posits that hydraulic geometries may estimate water depth in unregulated streams, under most flow conditions. The ability to estimate depth by these means is valuable where pre-collected data is minimal and highly accurate depth-discharge relationships are not required. Future work will examine three important points: 1) if the observed relationship between depth and hydraulic geometries is consistent in other regions of the world; 2) if hydraulic geometries are useful for calculating depth-discharge rating curves; 3) can hydraulic geometries be useful as a standalone mechanism for large scale inundation estimation.

Research Question 2

Flood Damage Wizard has proven that it can make use of a variety of Address Point-Cadastre datasets and generate feasible outputs. *Flood Damage Wizard* utilizes a fuzzy logic based, text matching algorithm. This algorithm is the Python module *FuzzyWuzzy*. In each of the three domains from this study, *FuzzyWuzzy* provides a suitable match between each Address Point-Cadastre site and depth-damage function. Further, more robust heuristic algorithms, such as expert systems or data mining approaches, such as Topic Modeling might offer advantages to

the use of simple fuzzy logic. The current process is iterative and as additional datasets will accompany additional refinement.

Results indicate that an approximate flood damage study is possible using the *Flood Damage Wizard* toolkit. How *Flood Damage Wizard* compares to other models, such as detailed Hazus analyses, and to real world scenarios will be a part of future study. However, this initial framework presents an opportunity to use a widely available dataset, which is typically not used in a dynamic nature. This usage would be akin to how the NFIE-Hydro system utilizes High Resolution Rapid Refresh (HRRR) datasets (NOAA, 2016).

The results of a *Flood Damage Wizard* run also begin to communicate the uncertainty of flood damage estimation through probabilistic prediction and use of uncertainty bounds in estimated age. *Flood Damage Wizard* assumes two extremes are possible – either the structure is brand new or completely depreciated. Additionally, all residential units have three depreciation functions applied to them. Using this methodology, if a user provides only a description of each building location, *Flood Damage Wizard* estimates a probabilistic distribution of at least three financial damage estimates and 15 economic damage estimates. Quantification of the uncertainty resulting from the use of default values will warrant additional investigation. *Flood Damage Wizard* could communicate the uncertainty in these estimates if both a mean and standard deviation are found for the default values. These methodologies require further investigation and serve to exemplify how to represent uncertainty within the model.

Considerable uncertainty exists when using the depth-damage functions. Freni et al. (2010) document this uncertainty in their study and suggest that better estimation of the depth-damage curves will reduce uncertainty. A method of comparison similar to that used by Apel et al. (2009) will provide guidance in how to investigate such uncertainty. Further, the address

points constitute approximate building locations. Both the curve utilized and the perceived location of the structure have considerable uncertainty associated with them. Attempts to quantify this uncertainty are an avenue of further research.

Optimization of the code within Flood Damage Wizard is necessary before achieving functionality in a real time environment. Using the local RC2 computing cluster at The University of Alabama, a 12-CPU simulation can take at least several hours, with larger flood events in highly developed areas taking several days. The cause of this bottleneck is the text matching processes occurring within the Python code. Thus, the system requires optimization before use in a real-time context. Optimization of the current fuzzy logic schema or the exploration of more robust tools, such as topic models or expert systems, are avenues of improving simulation speed. Additionally, a lookup table which matches common parcel descriptions to the best depth-damage function, prior to implementation of the fuzzy logic algorithm, may also expedite processing.

Research Question 3

This research considers if rapid, high resolution, large scale flood damage modeling for buildings can approximate total losses comparable to detailed regional studies. This approach utilizes the descriptions of property provided by cadastral datasets assembled by county tax assessors and performs a fuzzy logic, best match analysis to determine the appropriate depth-damage relationship for each property. The *Flood Damage Wizard* toolkit provides this functionality. The researchers consider which inputs are significant for the model to provide results that closely match those from a detailed Hazus Level II analysis. Results indicate that market value, use description, and square footage, are optimal for the fuzzy logic, utilizing

cadastral datasets, approach to provide total damage results that are comparable to Hazus Level II datasets. If the parcel data contains vague use descriptions, the user must exercise caution. Despite the similar inconsistencies in classifying building occupancy or use found in Hazus Level II inventory and cadastral datasets, the results have shown that this method can provide a quick real-time regional large scale flood damage estimate within a short lead time. Such damage estimates can provide valuable information in the forecast of regional economic impacts and thus help decision makers in the management of a flood event response.

To avoid mismatches between parcel descriptions and depth-damage descriptions, it may benefit the researchers to consider a weighted input scheme. For example, consider the “Small Grocery” curve description paired with “Small Tract” parcel description in Table 18. The word “Small” has very little influence on the proper match in either the parcel or curve description. The inputs “Grocery” and “Tract” are much more significant in determining a proper match. The current text matching schema matches by assigning equal weights to all terms in a particular text string. A method by which to weight inputs were “Small” is weighted less than either “Grocery” or “Tract” would likely produce a much better text match. *Flood Damage Wizard* assigns a match value to each text string match and will disregard any matches that do not meet a minimum match value (Gutenson et al., 2016). Such a technique could lower the match value score assigned in cases such as this. A lower match value score would then be ignored by Flood Damage Wizard.

General Conclusions

This work serves to illustrate that a rapid flood damage system, utilizing national hydrologic forecasts, is achievable in nearly every section of the CONUS. Locations in which stream flows are heavily regulated and anthropogenic changes to channel geometries occur are

likely to have data requisite for detailed study available. Also, a robust building inventory for flood damage calculation is likely available in such developed locations. For example, the large city of Austin, Texas has a robust flood early warning system with over 200 HEC-RAS models estimating flood depths (City of Austin, Flood Early Warning System, Accessed January 21, 2016, <https://www.austintexas.gov/department/flood-early-warning-system>). This research may work to fill the data gaps where no thorough ground survey of environmental and/or economic characteristics exists. These locations are likely rural and underserved by the establishment.

This research demonstrates that hydraulic geometries can approximate channel geometries for unregulated streamflows where no human alterations to the channel exist. This work supports the discussion introduced by Hodges (2013) who proposes several methods, including hydraulic geometries, to develop synthetic cross sections for continental scale river dynamics simulation. These channel geometries may be combined with existing, nationwide 10-meter DEM datasets to develop synthetic cross sections with approximate flood plain information in areas where no detailed survey of cross sections is available. These cross sections would serve as an input into a one-dimensional hydraulic model. Existing models, such as Simulation Program for River Networks (SPRNT), are available which can utilize such information for efficient hydraulic modeling (Liu and Hodges, 2014). For mapping and damage assessment purposes during flood events, the hydraulic model output of interest is depth.

Alternatively, the research demonstrates that an approximate depth estimate can be made for high flow events using hydraulic geometries available in the East Gulf Coastal Plain of Alabama. These hydraulic geometry depth estimates may serve to approximate hydraulic depth of streams with limited survey information. For example, utilizing 10-meter DEMs, NHDPlus Version 2 flowlines, and forecasted flows from the National Water Model, it may be possible to

approximate, rapidly, flood depths with the use of only the hydraulic geometry depth coefficients in the East Gulf Coastal Plain. This finding is unique, in that little existing research examines the use of hydraulic geometries in non-bankfull discharge scenarios.

Once an approximate means of estimating flood water depths is made available, the researcher can estimate resulting damage to buildings. This research demonstrates the fuzzy logic platform, *Flood Damage Wizard* can approximate a detailed flood damage survey. The researcher accomplishes the approximate flood damage survey by utilizing cadastral data with approximate building locations (Address Points), clear use descriptions, market value estimates, and square footage. As with the use of hydraulic geometries, this method is not a substitution for detailed study. However, *Flood Damage Wizard* has potential to approximate flood damage in data poor environments.

With these concepts, the potential to understand both the environmental and socioeconomic impacts of a predicted flood increase for all of the CONUS. This knowledge can expedite decision maker's ability to better understand the short- and long-term economic impacts of floods, provide a means to reduce flood impact in the short term, and expedite the process of recovering and learning from flood events.

However, this knowledge pairs with an understanding that these estimates are approximations, with associated uncertainty. Though all results of this research demonstrate significance, existing uncertainty pervade the results. For example, the use of a stochastic model (hydraulic geometry) for inundation calculation and channel approximation is subject to random error and local conditions. The use of these methodologies must occur in situations where the user anticipates considerable uncertainty.

Additional Research

If the tools identified here are of future value to decision makers, researchers must determine if the answers given by the solutions are good enough. As with all models, outputs are abstractions of reality and their usefulness ties to where and how they best approximate reality. A lack of real world measurement in floods and resulting damage precludes the researcher from examining how well these models approximate the real world. Comparison to actual events is a topic of further research.

As the global society advances in the Information Age, big data will progress to provide additional resources for large scale modeling at localized resolutions. This evolution will inevitably influence how researchers look at modeling and its possibilities. The improvements in hydrologic modeling capabilities spur forward a new way to look at modeling flood inundation and flood damage in short term time frames. As new data becomes available, how it can augment what this research proposes will change accordingly. This new information may come from passive observation (remote sensing), database aggregation and enhancement, or collaboration between private and public institutions. As with the goals of the OWDI and in general the U.S. Government's Open Data Initiative, providing data streams to researchers is of great importance to provide easy access to such data for efficient utilization.

An alternative means of generating channel estimations is through remote sensing applications. Mersel et al. (2013) discuss the use of the Surface Water and Ocean Topography (SWOT) satellite (planned to be launched in 2020) in estimating water topography. Yoon et al. (2012) also explore a method of using SWOT data in estimating river bathymetry. These studies suggest that the use of SWOT in estimating unknown bathymetry is promising.

The use of LiDAR-based analysis is another venue for channel geometry estimations. Hilldale and Raff (2007) evaluate LiDAR's effectiveness and find that the results display consistent bias. Kinzel et al. (2013) note that this form of LiDAR is still in the developmental/exploratory phase of usage. Legleiter et al. (2015) note that the bathymetric LiDAR utilized in their study was unable to detect shallow depths and accuracy was highly reduced in turbid streams. However, the prospect of using Space-Borne green-wavelength LiDAR to map bathymetry is promising though and will be of great interest in future studies (Adhallah et al. 2013). Using this LiDAR-based process could result in much more accurate channel representations. Continued development in this field will be critical in riverine hydraulic model performance.

Though the fuzzy text matching algorithms employed here provides relevant results, topic models may provide a more robust text matching schema. Topic Models are a data mining approach to interpreting the relationships between strings, usually in the form of structured documents. In topic modelling, a collection of documents or text strings form a collective corpus. Topic modeling assumes that the theme or themes of a document or string closely align to the words that comprise the document or string. Thus, instead of searching for information using a keyword approach, data exploration occurs through a themed approach. Additionally, multiple themes can and are typically captured in one document (Blei 2012).

Topic Models can mine social media (Lim et al., 2013; Paul and Dredze, 2014) and analyze web supplied user reviews (Rossetti 2015). Essentially, the use of topic models pervade any repository in which information overload is a possibility (Yang et al., 2015). Traditionally, LDA Topic Models categorize structured strings or documents and do not perform well in short, unstructured environments (Zhao et al., 2011; Lim et al., 2013). However, Lim et al. (2013) and

Zhao et al. (2011) both prove that they are usable in unstructured environments as well. The researcher envisions that Topic Models may assume two roles in furthering this research. The first is as a robust means of matching a parcel description with the appropriate depth-damage function. Secondly, the researcher envisions this tool's use as an alternative to depth-damage functions, in a manner akin to that used by Merz et al. (2013). Access to detailed flood insurance claims might supply a corpus for examination by a Topic Model. A flood damage Topic Model may determine how local flood conditions ultimately impact particular structures.

An alternative to matching depth-damage functions to parcels through fuzzy relationships is developing a standardized parcel dataset. A standardized dataset would provide a means of automatically assigning a given depth-damage function to a given parcel description because a standardize format implies predictability in the dataset. Both the Committee on Land Parcel Databases (2007) and Abt Associates Inc. (2013) highlight investigations into a standardized national dataset; none exists in the contemporary. Collaboration between these efforts and those ongoing through the OWDI may bring additional support for a nationalized parcel database which is capable of supporting natural hazard impact modeling.

Many video game connoisseurs know the video game *SimCity* (Electronic Arts, *SimCity*, Accessed January 22, 2016, <http://www.simcity.com/>). In this virtual environment, players build a city and manage social and governmental dynamics of the city as the cities primary decision maker (Woessner, 2015). The researcher envisions that the work which composes this dissertation may be useful as a component of a virtual interplay between social and environmental dynamics with a cloud-based, *SimCity* front end. Hartevelde (2008) proposes a flood decision support system with a *SimCity* front end. With the buildup of additional data streams, both social and physical, such a framework will become feasible. For example, the

continued use of BIM models may allow for the development of virtual, three-dimensional cities. Existing models can describe the interactions among social and physical systems. Additional environmental hazards (earthquakes, drought, etc.) and processes (permitted pollutant discharges, etc.) may also be interwoven in this scheme. Decision makers at different levels of local government may be able to act in their virtual positions within the *SimCity* framework. Multiple municipalities may also interact with one another. The decision maker's actions can affect the scenarios presented to one another. Borgdorff et al. (2015) design a system with a similar architecture for urban decision support. Such a technology is promising as an avenue of future research. The term *serious gaming* describes this concept of applying video games to function as more than a form of entertainment (Ritterfeld et al., 2009). The USGS experiments with serious game usage now (Wein and Labiosa, 2013). A quasi-*SimCity serious game* offers the compelling ability to educate and train decision makers, may foster cooperation between multiple municipalities, and might also promote advanced decision making with little training overhead.

REFERENCES

- Abdallah, H., J. Bailly, N. N. Baghdadi, N. Saint-Geours, and F. Fabre. (2013). Potential of space-borne LiDAR sensors for global bathymetry in coastal and inland waters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(1), 202-216. DOI: 10.1109/JSTARS.2012.2209864
- Abt Associates Inc. (2013). The feasibility of developing a national parcel database: County data records project final report. Bethesda, MD: U.S. Department of Housing and Urban Development.
- Apel, H., G. Aronica, H. Kreibich, and A. Thielen. (2009). Flood risk analyses—how detailed do we need to be? *Natural Hazards*, 49(1), 79-98. DOI: 10.1007/s11069-008-9277-8
- Archfield, S. A., M. Clark, B. Arheimer, L. E. Hay, H. McMillan, J. E. Kiang, et al. (2015). Accelerating advances in continental domain hydrologic modeling. *Water Resources Research*. DOI: 10.1002/2015WR017498
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84. DOI: 10.1145/2133806.2133826
- Borgdorff, J., H. Krishna, and M. H. Lees. (2015). SIM-CITY: An e-Science Framework for Urban Assisted Decision Support. *Procedia Computer Science*, 51, 2327-2336. DOI: 10.1016/j.procs.2015.05.399
- Committee on Land Parcel Databases. (2007). National land parcel data: A vision for the future. Washington, D.C.: National Academy Press.
- Das, S. and R. Lee. (1988). A Nontraditional Methodology for Flood Stage-Damage Calculations. *Journal of the American Water Resources Association (JAWRA)*, 24(6), 1263-1272.
- Ding, A., J. F. White, P. W. Ullman, and A. O. Fashokun. (2008). Evaluation of HAZUS-MH flood model with local data and other program. *Natural Hazards Review*, 9(1), 20-28. DOI: 10.1061/(ASCE)1527-6988(2008)9:1(20)
- Eckstein, O. (1965). *Water Resource Development – The Economics of Project Evaluation*. Harvard University Press, Cambridge, Massachusetts.

- Federal Emergency Management Agency (FEMA). (2014). Risk MAP flood risk products. Washington, D.C.: Federal Emergency Management Agency (FEMA).
- Freni, G., G. La Loggia, and V. Notaro. (2010). Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation. *Water Science and Technology*, 61(12), 2979. DOI: 10.2166/wst.2010.177
- Grigg, N. S. and O. J. Helweg. (1975). State-of-the-Art of Estimating Flood Damage in Urban Areas. *Journal of the American Water Resources Association (JAWRA)*, 11(2), 379-390.
- Gutenson, J. L., Ernest, A. N. S., Oubeidillah, A. A., Zhu., L., and Sadeghi, S. T. (2016). Rapid flood damage assessment using public domain cadastral and address point data with fuzzy logic algorithms. *Journal of the American Water Resources Association (JAWRA)*. Manuscript submitted for publication.
- Harteveld, C. (2008). A Playful Approach to Flood Defence. In: *4th International Symposium on Flood Defence*. Toronto, Ontario, Canada.
- Hilldale, R. C. and D. Raff. (2008). Assessing the ability of airborne LiDAR to map river bathymetry. *Earth Surface Processes and Landforms*, 33(5), 773-783. DOI: 10.1002/esp.1575
- Hodges, B. R. (2013). Challenges in continental river dynamics. *Environmental Modelling & Software*, 50, 16-20. DOI: 10.1016/j.envsoft.2013.08.010
- Kinzel, P. J., C. J. Legleiter, and J. M. Nelson. (2013). Mapping River Bathymetry With a Small Footprint Green LiDAR: Applications and Challenges1. *Journal of the American Water Resources Association (JAWRA)*, 49(1), 183-204. DOI: 10.1111/jawr.12008
- Legleiter, C. J., B. T. Overstreet, C. L. Glennie, Z. Pan, J. C. Fernandez-Diaz, and A. Singhanian. (2015). Evaluating the capabilities of the CASI hyperspectral imaging system and Aquarius bathymetric LiDAR for measuring channel morphology in two distinct river environments. *Earth Surface Processes and Landforms*. DOI: 10.1002/esp.3794
- Leopold, L. B. and T. Maddock Jr. (1953). The Hydraulic Geometry of Stream Channels and Some Physiographic Implications. U.S. Geological Survey Professional Paper 252, 57 pp. <http://pubs.er.usgs.gov/publication/pp252>
- Lim, K. W., C. Chen, and W. Buntine. (2013). Twitter-network topic model: A full Bayesian treatment for social network and text modeling. In: *NIPS2013 Topic Model Workshop*, pp. 4.
- Liu, F. and B. R. Hodges. (2014). Applying microprocessor analysis methods to river network modelling. *Environmental Modelling & Software*, 52, 234-252. DOI: 10.1016/j.envsoft.2013.09.013

- Maidment, D. R. (2014). National Flood Interoperability Experiment. *2014 American Geophysical Union (AGU) Fall Meeting*. American Geophysical Union, San Francisco, California.
- Merwade, V., A. Cook, and J. Coonrod. (2008). GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. *Environmental Modelling & Software*, 23(10), 1300-1311. DOI: 10.1016/j.envsoft.2008.03.005
- Merz, B., H. Kreibich, and U. Lall. (2013). Multi-variate flood damage assessment: a tree-based data-mining approach. *Natural Hazards and Earth System Science*, 13(1), 53-64. DOI: 10.5194/nhess-13-53-2013
- National Oceanic and Atmospheric Administration (NOAA). (2016). The high-resolution rapid refresh (HRRR). Retrieved from <http://ruc.noaa.gov/hrrr/>
- Oliveri, E. and M. Santoro. (2000). Estimation of urban structural flood damages: the case study of Palermo. *Urban Water*, 2(3), 223-234. DOI: 10.1016/S1462-0758(00)00062-5
- Oubeidillah, A. A., S. Kao, M. Ashfaq, B. S. Naz, and G. Tootle. (2014). A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US. *Hydrology and Earth System Sciences*, 18(1), 67-84. DOI: 10.5194/hess-18-67-2014
- Pappenberger, F., H. L. Cloke, D. J. Parker, F. Wetterhall, D. S. Richardson, and J. Thielen. (2015). The monetary benefit of early flood warnings in Europe. *Environmental Science & Policy*, 51, 278-291. DOI: 10.1016/j.envsci.2015.04.016
- Paul, M. J. and M. Dredze. (2014). Discovering health topics in social media using topic models. *PLoS ONE*, 9(8). DOI: 10.1371/journal.pone.0103408
- Ritterfeld, U., M. Cody, and P. Vorderer. (2009). *Serious games: Mechanisms and effects*. Routledge, New York, New York.
- Rossetti, M., F. Stella, L. Cao, and M. Zanker. (2015). Analysing User Reviews in Tourism with Topic Models. In: *Information and Communication Technologies in Tourism 2015*, pp. 47-58
- Sood, A. and V. Smakhtin. (2015). Global hydrological models: a review. *Hydrological Sciences Journal*, 60(4), 549-565. DOI: 10.1080/02626667.2014.950580
- Tate, E., C. Muñoz, and J. Suchan. (2015). Uncertainty and Sensitivity Analysis of the HAZUS-MH Flood Model. *Natural Hazards Review*, 16(3), 04014030, DOI: 10.1061/(ASCE)NH.1527-6996.0000167.
- Wein, A. and W. Labiosa. (2013). Serious games experiment toward agent-based simulation. Open File Report 2013-1152, 30 pp.

https://www.researchgate.net/publication/259078831_Serious_Games_Experiment_toward_Agent-based_Simulation

- Woessner, M. (2015). Teaching with SimCity: Using Sophisticated Gaming Simulations to Teach Concepts in Introductory American Government. *PS: Political Science & Politics*, 48(2), 358-363. DOI: <http://dx.doi.org.libdata.lib.ua.edu/10.1017/S104909651400211X>
- Yang, C. and C. Tsai. (2000). Development of a GIS-Based Flood Information System for Floodplain Modeling and Damage Calculation. *Journal of the American Water Resources Association (JAWRA)*, 36(3), 567-577. DOI: 10.1111/j.1752-1688.2000.tb04287.x
- Yang, G., D. Wen, N. Chen, and E. Sutinen. (2015). A novel contextual topic model for multi-document summarization. *Expert Systems with Applications*, 42(3), 1340-1352. DOI: 10.1016/j.eswa.2014.09.015
- Yoon, Y., M. Durand, C. J. Merry, E. A. Clark, K. M. Andreadis, and D. E. Alsdorf. (2012). Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of Hydrology*, 464, 363-375. DOI: 10.1016/j.jhydrol.2012.07.028
- Zhao, W. X., J. Jiang, J. Weng, J. He, E. Lim, H. Yan, et al. (2011). Comparing twitter and traditional media using topic models. In: *33rd European Conference on IR Research*, pp. 338-349.

APPENDICES

Flood Damage Wizard Code

The following text highlights how each module of *Flood Damage Wizard* functions. The code is written in Python 2.7. The modules which comprise *Flood Damage Wizard* are `Extract_Raster_Value`, `Damage_Curve_Search`, `Data_Merger`, and `Flood_Damage_Calculator`. *Flood Damage Wizard* requires two datasets, a building point shapefile and an inundation raster. These spatial datasets must have the same coordinate or projection system which describes their geographic location.

Extract_Raster_Value

This module uses the point shapefile of building locations to extract the depth value in the inundation raster at each point location. The analysis also identifies what inputs are within the point shapefile. The inputs basement presence and number of stories append to the use description to enhance the text description of each building. For instance, if the use description is “single family residential” and the data indicates that a basement is present and the residence is two stories, the use description becomes “single family residential basement two stories level”. Each additional quantitative parameter of a building such as square footage, market value, and age of structure do not append to the text description.

Damage_Curve_Search

Using fuzzy text matching, this module matches depth-damage functions to the use description provided by the building point shapefile. This use description may or may not be enhanced by the additional basement or number of stories descriptors. The singular best match or the one with the highest match value for each building is chosen based on the fuzzy text matching algorithm. The match must attain a match value of at least 60 to pass through as a valid match. Only one match is made between a building and depth-damage function. If no suitable match is found for the building, no damage estimate is provided. The assignment of a depth-damage function to a building also designates its occupancy which allows for the assignment of default values from Table 3 to fill data gaps. Matches to both structural, content, and inventory depth-damage functions are made for each building.

Data_Merger

This code examines the inundated building locations and determines what inputs have been provided and which are needed for further analysis. It uses the Table 3 default parameters to determine fill any data gaps. There are additional functions in this module which reassign damage outputs to the building point shapefile.

Flood_Damage_Calculator

Using what inputs the building point shapefile provides and any default values, this module of *Flood Damage Wizard* calculates both full replacement or financial damage and economic or depreciated damage. When market value data is within the building point shapefile, this along with full replacement value calculate both financial and economic damage. When no market

value is available, *Flood Damage Wizard* calculates a depreciated full replacement value which is a product of an occupancy related depreciation functions and structural age. These curves estimate depreciation based upon the structures age. A probabilistic distribution of both economic and financial damages are estimated. This provides multiple damage estimates for each of the building locations. These distributions of damage can aggregate as the average of the distribution or as the maximum value of the distribution. Estimated structural, content, and inventory estimates comprise total damage or loss to a structure.