ASSESSING THE RESOLUTION EFFECTS OF DIGITAL ELEVATION MODELS
ON AUTOMATED FLOODPLAIN DELINEATION
A CASE STUDY FROM THE CAMP CREEK WATERSHED IN MISSOURI

_____

A Thesis presented to

the Faculty of the Geography Program

University of Missouri-Columbia

_____

In Partial Fulfillment

of the Requirements for the Degree

Masters of Arts

_____

RICHARD CHARRIER

Dr. Yingkui (Philip) Li
Thesis Supervisor

December 2009

The undersigned, appointed by the Dean of the Graduate School,

have examined the thesis entitled:

ASSESSING THE RESOLUTION EFFECTS OF DIGITAL ELEVATION MODELS ON AUTOMATED FLOODPLAIN DELINEATION
A CASE STUDY FROM THE CAMP CREEK WATERSHED, MISSOURI

Presented by Richard J. Charrier

A candidate for the degree of Master of Arts

And hereby certify that in their opinion it is worthy of acceptance.

_____

Dr. Yingkui (Philip) Li
Assistant Professor
Department of Geography

_____

Dr. C. Mark Cowell
Associate Professor
Department of Geography

_____

Dr. Kathleen Trauth
Associate Professor
Civil and Environmental Engineering

_____

Dr. Soren Larsen
Assistant Professor
Department of Geography

Acknowledgements

I would like to thank all the people who helped and inspired me during this Master thesis work. I would especially like to thank Dr. Yingkui Li for his help and inspiration. His cheerful yet focused demeanor made this process enjoyable for all. His willingness to help, offer advice, and constant availability made all this work possible.

I would also like to thank Dr. Soren Larsen, Dr. Mark Cowell, and Dr. Kate Trauth for their help and guidance as members of my thesis committee. Dr. Larsen helped tremendously with his suggestions during the writing process. Dr. Trauth gave me opportunity to present my thesis results, directing me with generous helpful outside advice from her engineering students.

I would also like to thank Dr. Chris Fulcher, Dr. Tony Prato, and Chris Barnett, who as Directors for the Center for Applied Research and Environmental Systems provided me with the opportunity to do the work needed for this project. They provided access to the CARES data, models, and equipment needed to do the analysis that made this work possible.

I would also like to thank the staff at CARES for providing me with help as needed throughout the process, especially Bryan Mayhan and Scott Burton who worked tirelessly to help modify models. Without their expertise, this project would not have been possible.

Most importantly, I would like to thank my wife and children for keeping me focused. Without my family none of this would be possible. They are the inspiration for all my work, both past and present.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AML   Arc Macro Language

APCA   Automated Proximity Conformity Analysis

BASINS   Better Assessment Science Integrating Point & Nonpoint Sources. This is a hydrology model developed by the US Environmental Protection Agency.

CARES   Center for Applied Research and Environmental Systems

DEM   Digital Elevation Model

DFIRM   Digital Flood Insurance Rate Map

DOQ   Digital Ortho-photo Quadrangle

DOQQ   Digital Ortho-photo Quarter Quadrangle

DTM   Digital Terrain Model

DRG   Digital Raster Graphic – Digital representation of USGS topographic maps

EPA   Environmental Protection Agency

FEMA   Federal Emergency Management Agency

FIRM   Flood Insurance Rate Map

GIS   Geographic Information Systems / Geographic Information Science

HEC-HMS   Hydrology Engineering Center - Hydrologic Modeling System. This is a hydrology model developed by the US Army Corps of Engineers.

HEC-RAS   Hydrology Engineering Center – River Analysis System. This is a hydraulic model developed by the US Army Corps of Engineers.

LiDAR   Light Detection and Ranging. This is a method of gathering surface elevations by 'shooting' the earth with Radar from airplanes.

NED   National Elevation Dataset

NFECT   Nebraska Flood Elevation Calculation Tool

NFIP   National Flood Insurance Program

SWMM   Storm Water Management Model. This is a hydraulic model
           developed by the US Environmental Protection Agency.

TIN   Triangular Irregular Network.

USGS   United States Geological Survey

WISE   Watershed Information SystEm. This is a hydrology model
           developed by the engineering firm Watershed Concepts.

WSPRO   Water Surface PROfile. This is a hydraulic model developed by the
           USGS.

Abstract

Automated floodplain modeling commonly requires Digital Elevation Models (DEMs) to represent the topography. As a raster representation of the Earth surface, changing a DEMs resolution (data cell size) has a profound impact on the floodplain delineation. Since 1995 DEM resolution has increased from 100- to 1-meter resolution. This thesis addresses how different DEM resolutions, and different DEM data sources, affect the outcome of modeled floodplain boundaries in the Camp Creek Watershed, a predominately agricultural watershed in Missouri. Two data sets are analyzed: a Light Detection and Ranging (LiDAR) terrain model re-sampled to 1-, 3-, 5-, 10-, 15-, and 30-meter resolutions and existing United States Geological Survey (USGS) 5-, 10-, and 30-meter DEMs. The floodplain delineation process includes hydrologic modeling, hydraulic modeling, and floodplain delineation. Each process includes various input parameters and outputs. Resultant stream networks, watershed boundaries, and floodplains are examined to evaluate the effects of different resolutions. Using 3- or 5- meter LiDAR DEMs produces data that agree with the 1-m data greater than the 90th percentile. The agreement also includes the 10-m DEM data when analyses remove the floodplain modeling cumulative discrepancy effects. Similar trends were not found when using the USGS counterparts; possibly due to the use of the same underlying source material to create the DEMs. When removing the cumulative distortion effect of resolution on the entire modeling process, LiDAR DEM floodplains displayed a 1-4% increase in goodness of fit. Analyzing the results of two separate hydraulic

models (HEC-RAS and CARES) finds little difference between their calculated flood surface elevations. Additionally, the thesis analyzes the data storage needs and processing time for modeling different resolutions, finding substantial savings in both as the underlying DEM resolution is decreased. The thesis begins to analyze how models are affected by input variables but many additional studies are needed. Further study of these variables is needed to determine if a single most appropriate model and DEM resolution exists, or what combination of models are appropriate for various types of automated floodplain modeling.

Chapter 1

Introduction

Water is essential to life on this planet but it can also destroy as it nourishes. Water-related natural disasters, such as flooding, hurricanes, cyclones, tsunamis, mudslides, and blizzards account for approximately 60% of all natural disasters world wide (Frech, 2006). Flooding is the largest natural disaster, accounting for almost 40% of all incidents, and it affects nearly all portions of the United States (US). Figure 1 provides monetary value of all natural disaster damages in the US for the last century as tracked by the Federal Emergency Management Agency (FEMA). While the data does not isolate flooding damage costs, it does show that spikes in damages occur in years with major flooding or hurricane events.

The damage to property is estimated at about $4 billion, along with approximately 200 deaths annually in the US alone (Pielke Jr. et al., 2002). As seen in Figure 1, about once every ten years there are spikes in damages due to specific mega-storms or events. World-wide, of the top ten worst natural disasters (as defined by loss of life) over the last 200 years; there were three floods, two cyclones, one tsunami, and four earthquakes. In 2007 the top ten natural disasters included four flooding events, one hurricane, and one cyclone (Time, 2008).

It was a surrealistic experience to drive along Highway 63 between Ashland and Jefferson City, Missouri during late July 1993. The brave souls that

ventured onto the roadway knew that the Missouri River was coming and there was nothing they could do if it arrived ahead of the peak prediction. Drivers watched anxiously for approximately an eight kilometer span as surges flowing down the river caused water to slowly lap at the driving surface.
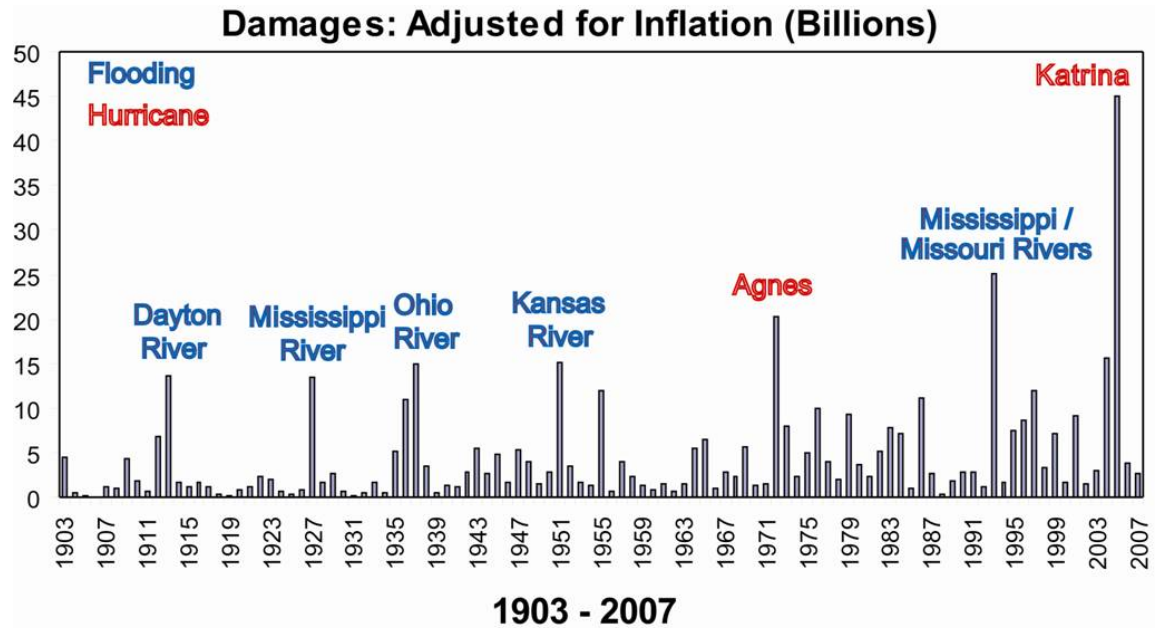
Figure 1: FEMA natural disaster damage estimates
Estimated damage includes all natural disaster types for the year (not just flooding)
Labels indicate major events that contributed to the increase in certain years
Data: http://www.fema.gov/library/index.jsp (April 2009)

The flood waters had risen over the previous month transforming the highway from a safe four-lane thoroughfare to a two-lane roadway before eventually being closed all together. In the following days, the junction of Highways 54 and 63 was almost completely under water with extensive flooding on both sides of the roadways. The bridge that spanned the normal Missouri River flow was dwarfed by the expanse of the flooding. By July 30, the Missouri River had reclaimed its entire floodplain (Figure 2).

Figure 2: Highway 54 / 63 interchange north of Jefferson City, MO
Looking south across flooding into Jefferson City, MO on July 30, 1993
Seen at the top of the floodplain is the bridge spanning the normal river channel
(Photograph from Missouri Highway and Transportation Department)

To an outside observer, the cars traveling along Highway 63 must have seemed to be floating along the water surface as no land surface could be discerned. The drivers glanced regularly to the north and saw bluffs rising directly from the river water. To the south, the only features to break the surface flow were the top of Renz Prison, a few business rooftops, and the branches from an occasional tree stand. While businesses were rebuilt after the flood waters receded, Renz Prison was abandoned as the extent of the devastation became apparent.

This scene was replayed along many roadways near the Missouri and Mississippi Rivers in 1993. Flooding can be quite extensive as was the case in 1993 in the Midwestern United States. Figure 3 shows a comparison of the Illinois, Missouri, and Mississippi River confluences near St. Louis, Missouri before and during the largest flooding event to hit the United States in the last 100 years. The "Great Flood of '93" lasted from May until September and caused record flooding across nine Midwestern states.

The "Great Flood of '93" caused numerous levee breaks resulting in fifty deaths, thousands of evacuations, and approximately fifteen billion dollars worth of damage (Larson, 1997). The flood of 1993 was termed a "500 year flooding event" and was followed two years later by the "100 year flooding event" in a large portion of the same areas hit during the floods of 1993. There was substantially less loss during the floods of 1995 because the area was still rebuilding from the record flooding two years prior. There was a second "500 year flooding event" along the Mississippi River in 2008 that caused many levee failures throughout Iowa and prompted many discussions about how there could be two "500 year floods" within 15 years of each other (Koster, 2008).

While this was an extreme case of riverine flooding, it represents a story told many times across the globe. During the flooding in 2008, nearly 1/3 of Iowa was declared a disaster area by President George W. Bush (AP, 2008), while less than a year later during March 2009 the majority of the State of North Dakota was deemed a disaster area by President Obama due to unprecedented flooding of the Missouri and Red Rivers (AP, 2009). Flooding of the Red River in

1997 prompted the City of Fargo to build a more extensive flood control system. While this levee system protected the city in 2009, levees generally confine flow resulting in pooled water upstream and faster and deeper water flow along the

Figure 3: St Louis, MO during the "Great Flood of '93"
Satellite imagery of a normal situation (1991) and "Great Flood of 1993"
(Photograph from NASA Earth Observatory)

constriction. In cases such as these, flood prevention measures are

implemented after a flooding event occurs while in other instances these

measures are implemented after models predict flooding events. Predictive

modeling of flood events is useful in identifying locations where measures that

can prevent loss of life and property should be implemented. These measures

don't prevent flooding rather they redirect water to less vulnerable locations.

There are many types of flooding events. While the scale of the flooding

event varies, the damage caused quite often can be avoided or at least

minimized. Riverine flooding is typically modeled to predict what is commonly

known as the "100 year" flooding event or the more proper term, the one percent

annual chance flood-hazard event. In most cases, riverine flooding is also

modeled for ten, fifty, and five-hundred year events. The numerical models can

only predict a flooding event based on its unique input parameters, hence a

comprehensive study of flood heights, locations, and tendencies are needed to

understand how floods behave.

Historically, flood maps were developed using complex equations relying

on elevations and discharge information derived from field data (Figure 4). In the

1960s, these calculations were incorporated into early versions of the US Army

Corps of Engineer's Hydrologic Engineering Center's (HEC) products to

determine flood water depths. Floodplain boundaries were then manually drawn

using USGS topographic maps. As geographic information systems (GIS) were

developed and made available, automation of the process became standard

(Noman et al., 2001).

The original process has the advantage of using field verified elevations and flow calculations and is still used within detailed delineations of flooding events. Automated floodplain delineation relies on derived calculations but eliminates the need for flood boundaries to be drawn manually. Additionally, field-based data collection can be costly and time consuming; consequently, this approach is limited to the most important areas of floodplain delineation.
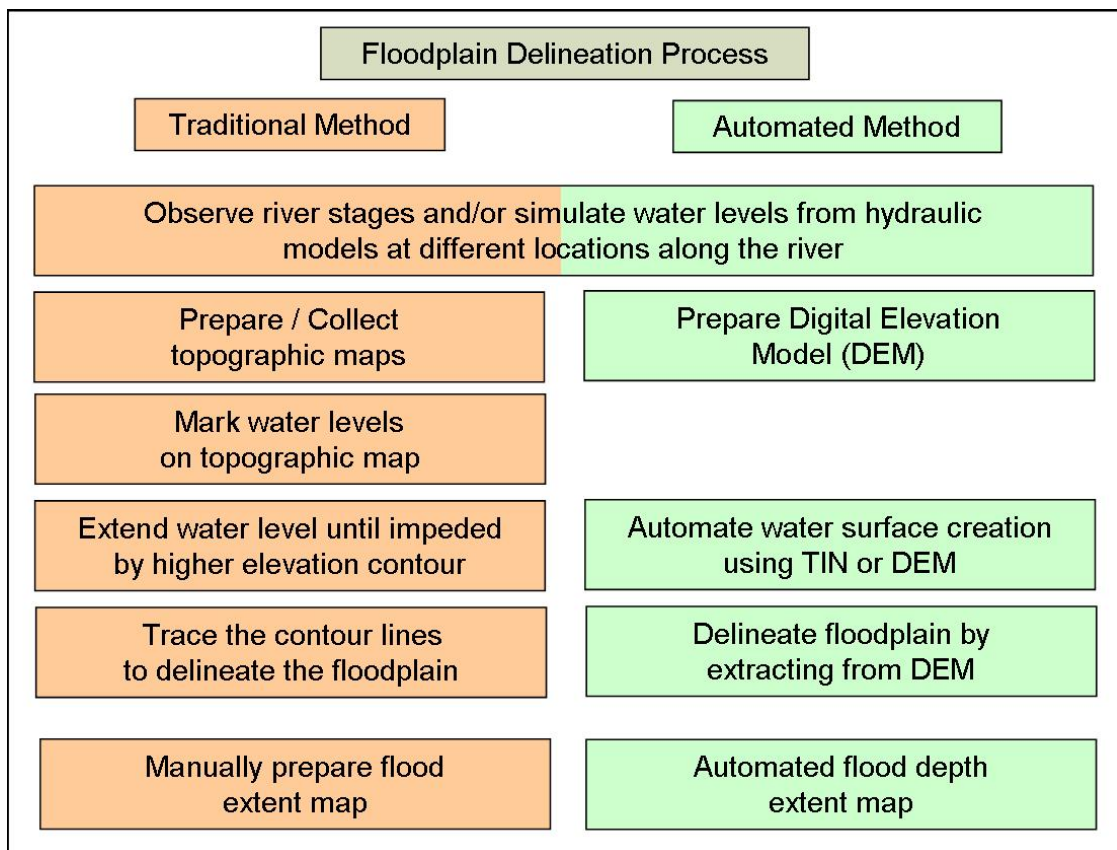
Figure 4: Historical floodplain delineation process
Two processes of delineating floodplains
Traditional methods are still used in some detailed studies

One important data source used in predictive modeling today is the digital elevation model (DEM). Known elevation points across the earth's surface are linked spatially to square cells organized into a grid. This digital representation of

the earth's surface is then used to replicate and visualize the properties of the terrain for different purposes. Using a grid to represent data is an older technology; most commonly associated with photographic images. The clarity of a photograph is based upon its resolution, or the size of its picture elements (pixels). Each pixel represents a single color so images with higher resolution allow for more shading and blending, resulting in a clearer and more accurate photograph. In the case of DEMs, each pixel holds a value of elevation for that pixel and a higher resolution results in a representation of the Earth's surface constructed of more individual elevations (Figure 5).
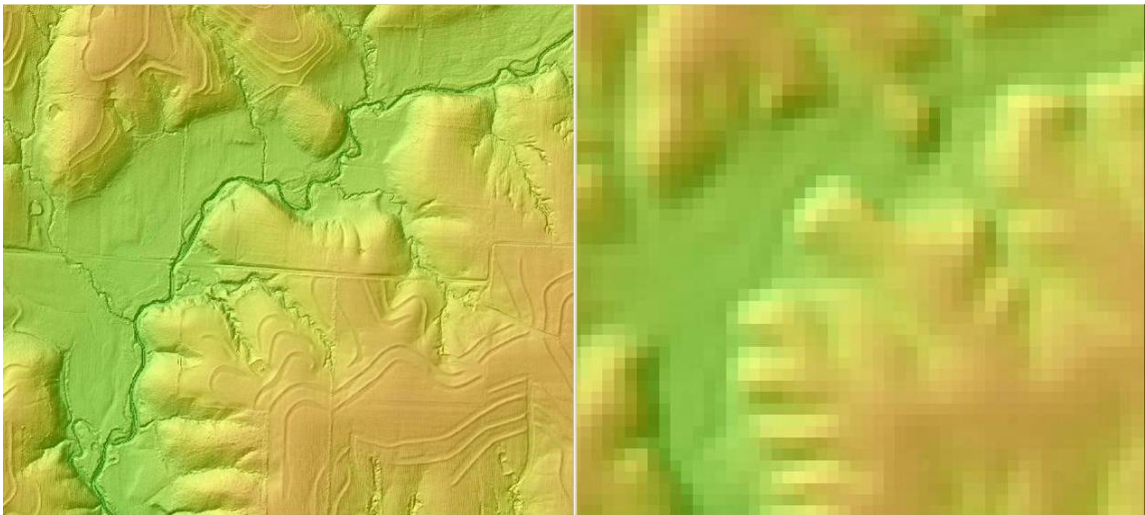
Figure 5: Comparison of a one-meter and thirty-meter DEM
DEM comparison of small portion of Camp Creek watershed

The technologies used to create DEMs have evolved over time. DEMs can be developed using digital aperture radar, digital image correlation, Light Detection And Ranging (LiDAR) data, or field surveys. A higher resolution DEM requires larger amounts of data storage and computer processing ability. During the past twenty years, computers have advanced in both processing capability

8

and storage capacity, allowing for the development and use of higher resolution DEMs. For example, the state of Missouri produced its first statewide thirty-meter DEM in 1998, with a ten-meter DEM following in 2005. Portions of Missouri have access to LiDAR one-meter DEMs, although a statewide dataset has yet to be developed. LiDAR is produced by using low flying airplanes to shoot laser pulses at the ground and then capture these on their return to the plane. By measuring the elapsed time between the transmission and reception of the pulse, an elevation is created for that location. By repeating this process millions of times, elevation data sets for large areas can be compiled.

The state of Missouri covers 180,545 square kilometers. The statewide one-hundred-meter DEM is approximately 125 megabytes (.125 gigabytes) in size. The same DEM at a ten-meter resolution is almost 15.3 gigabytes. The one-meter LiDAR DEM used for this study was 36 megabytes that would extrapolate to approximately 140 gigabytes for an area the size of Missouri. Along with the added processing requirements, the storage needs of a higher resolution DEM are typically offset by the more accurate elevation representations.

The availability of digital representations of the Earth's surface allows for multiple types of numerical modeling. DEMs have become base data for a range of models, the most prevalent of which are models that track or predict water movement (hydrologic models). Along with hydrologic models, DEMs have many other applications that include lava flow analysis (Szekely and Karatson, 2004), landslide (Stolz and Huggel, 2008), species (Guisan et al., 2005) and vegetation

(Hawkins et al., 2007) distributions, and glacial modeling (Berthier and Tourtin, 2007), to name only a few. Some models use DEMs (higher or lower resolution) as a focal point of the analysis, while other models only use DEMs in a peripheral manner. For this reason, it is essential to study the effect of DEM resolution on model results and determine the optimal resolution for a certain type of model and application so that the resulting information is accurate.

Floodplain modeling is one field of study that relies on DEMs as base data. Modeling floodplains requires, at a minimum, a combination of hydrologic (discharge), hydraulic (flood elevation), and delineation (flood location) models. The combination of these three types of models produces the information needed to delineate floodplains. While other models can also be used to delineate floodplains, they all rely on the most accurate information about the terrain of the target area. Knowing the extent of floodplains is vital to disaster and emergency management and preparation. In addition, modeling of floods is also required during actual disasters to help emergency responders.

The different portions of floodplain modeling all need DEMs as base data. There is a general assumption that higher resolution DEMs will produce more accurate results. LiDAR data can produce DEMs accurate to better than one-meter resolution. The most common DEM resolutions used for hydraulic models and floodplain delineations are of much lower quality than LiDAR provides (Omer et al., 2003). Currently, USGS provides a seamless thirty-meter DEM for the entire country. On a local scale the USGS also provides higher resolution data. For example, they provide a ten-meter state-wide DEM for Missouri and five-

meter localized watershed DEMs. These DEMs were created from USGS

1:24,000 scale topographic data, while the LiDAR data has been developed

within only the past several years.

The nature of an actual flooding event is quite literally fluid. Predictive

flooding models assign a boundary for a floodplain and claim it as the peak

discharge for a flooding event at a given return interval. While a predictive model

can assume normal conditions, slope, and velocity of a floodplain during the

event, the surface of the water itself does not remain constant for reasons that

include wind as well as natural and man-made debris. Flooding events can also

vary within larger watersheds. For example, some of a watershed's tributaries

may experience only enough precipitation for a ten percent annual chance flood

event while another portion might be inundated with enough water to create a 0.2

percent chance flooding event. There are also many types of flooding events

beyond normal riverine flooding such as flash flooding or storm surges. There

are also man-made flood prevention measures that can alter natural flooding or

fail altogether, causing additional damage (Figure 6).

Figure 6: Canal levee failure in New Orleans, LA
Example of man-made flood protection measure failure increasing damage
Shown is the 17th Street canal that failed during hurricane Katrina in 2005

1.1 Thesis Objectives

The main purpose of this thesis is to contribute to the goal of creating the most accurate automate floodplain models possible. To do that it is important to understand how many potential variables affect the automated floodplain modeling process. This thesis is designed to assess the impact of DEM resolution on automated floodplain modeling. This is examined by analyzing delineated stream networks (flow paths), delineated drainage basins, and modeled floodplains that are derived using the same process but with varying underlying DEM data.

The resolution of the underlying DEM used for analysis affects all portions of the floodplain modeling process and is the one variable this thesis isolates for study. As DEM resolution is one of the more important variables of the automated floodplain delineation process, this thesis isolates this by keeping all other variables constant. Additionally the study does not attempt to create a flooding event based on any actual flood, but rather attempts to create nine idealized flooding events that are compared to each other for the sole purpose of isolating the effect of DEM resolution.

Since the automated floodplain modeling process is comprised of three distinct modeling processes, it stands to reason that the variations between results will compound at each step in the process. Variations in data results from resolution will have a cumulative affect on the floodplain that ultimately delineated. This thesis isolates the data to quantify the amount of cumulative variation that is introduced by using a standardized set of water surface elevations as a secondary analysis.

There were two distinct sets of DEM data that are available for the study area. This thesis also identifies the amount of correlation that exists between DEM resolution and the original data sources of the DEMs. Isolating the computed results and comparing them between the two sources of DEM data provides an insight into the way that the resolution of a DEM is affected by its source data, and ultimately how those variations affect the modeled floodplain that is produced.

Finally, this study is designed to provide future studies with a baseline for analysis of other automated floodplain model variables. DEM resolution is one of many types of variables that affect a predicted flooding event and all of these variables need to be studied to ultimately determine the best model choices. Prevention of both property damage and loss of life is the goal of any floodplain model and this thesis is designed to work with other studies to find a best possible method or model to accomplish this goal.

Chapter 2

Literature Review

2.1 DEMs and LiDAR

Many floodplain delineation models use DEMs as the best available source for terrain modeling. A DEM is a raster dataset and its resolution is determined by the cell size. As technology has improved, the storage capacity of the datasets has increased exponentially, allowing DEMs to be created at much higher resolutions. A higher resolution DEM offers a more accurate portrayal of the Earth's surface. For example, LiDAR can collect enough information to support the creation of DEMs at a resolution of less than one meter, although such data demands a large digital storage space and computer processing capability (Tate et al., 2002). The technological development and application of LiDAR allows for the ability to capture more accurate terrain information and enhanced modeling procedures resulting in more accurate floodplains.

The varying sources of data used to create DEMs should be studied to see how they impact the accuracy of the DEM. Resampling a one-meter LiDAR DEM to different resolutions allows for the examination of differences due to variations in DEM creation techniques prior to modeling, and it also accounts for some statistical variance prior to the final analysis. Collecting various DEMs in addition to resampling provides a good range for studying resolutions, allowing for variations of source data. The most accurate DEMs are derived from LiDAR

data because this represents a multitude of data points over even the smallest areas, creating intricate raw elevation data.

Using LiDAR data to create a DEM, however, requires the developer to remove points of false elevation gathered during data capture. LiDAR itself can also be manipulated to create a more-or-less accurate DEM through factors during data collection, filtering, and cleanup (Noman et al., 2003). As the LiDAR raw data is collected there are many millions of data points analyzed. Each data collection point represents a series of data returns with the last elevation returning to the origin interpreted as the lowest elevation and therefore represents the earth's surface. The filtering process removes all but the last returns to transform a data set with millions of elevation points to one with only those believed to be the 'bare earth' surface (Figure 7).
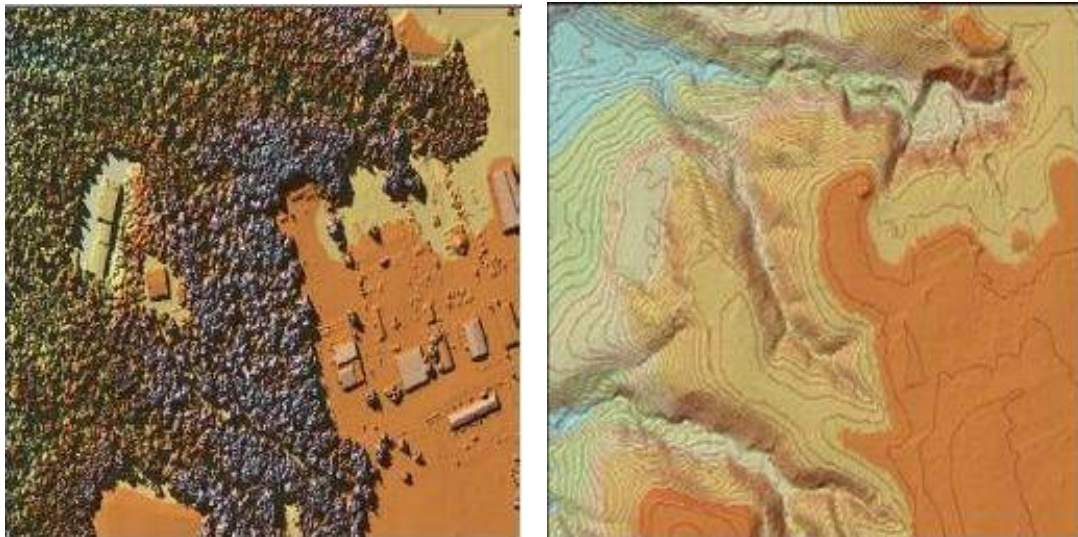
Figure 7: LiDAR first return data and bare earth data
Initial LiDAR data before and after being stripped of vegetation and buildings
Imagery: July 22, 2008 MGISAC LiDAR White Paper #1

Omer et al. (2003) suggested that using all data points from LiDAR raw data would create a cumbersome, and in some places redundant, dataset. Filtering raw LiDAR data can remove excess and redundant points without reducing resolution. That study tested different filtering methods and found that a filter of up to four degrees resulted in negligible accuracy disruption in the final floodplains delineated. The analysis methodology for that study looked at the flooding area percent differences, a useful analysis when comparing large area polygons. Omer et al. (2003) also describes the complications with using LiDAR data sets because of their size and computational requirements, and also studied how different filtering techniques affected the processing time. The study states that using a four degree filter provides an average processing time at 55% of the original data processing of a one degree filter.

Since a higher resolution DEM is a more accurate representation of the Earth's surface, it stands to reason that using a higher resolution DEM should result in more accurate floodplain boundaries. While it is known how the variations of resolution affect the cells of an individual DEM, DEMs are used for many different types of analysis beyond flooding models (Gallant & Hutchinson, 1997). Each type of modeling determines the resolution that best works for each particular type of study. The higher the resolution of the DEM, the more accurate the Earth's surface is represented, although this result should be studied to understand if the assumptions about DEM resolution for floodplains holds true.

Along with different resolutions, there are other errors that can be introduced into DEMs, and their effects have to be accounted for with regard to

flood modeling. Because a DEM is a raster dataset, the cell size is not the only variant. In using different resolution DEM data, it is important to be consistent while delineating the study area. To account for low-lying areas, DEMs must be filled to ensure water flow paths are uninterrupted, and for LiDAR DEMs hydrological enforcement using the same stream network must be undertaken to account for flow paths through relatively flat areas. Both of these steps can introduce uncertainties into the DEMs and require careful application. There have been studies about how these actions alter the DEMs with regard to terrain analysis (Holmes et al., 2000, Yilmaz et al., 2004, Callow et al., 2006) finding that no single method for DEM correction performs the best for all applications.

Comparison analysis of a LiDAR DEM against a thirty-meter USGS DEM with regard to a known flooding event also proves useful when studying how DEM resolution affects floodplains. A study of flooding events in Pitt County, North Carolina (Wang & Zheng, 2005) analyzed the different predicted extents produced by a thirty-meter USGS National Elevation Dataset (NED) DEM and a one-meter LiDAR DEM. The study itself did not analyze the differences created by using different resolution DEMs, but focused instead on the differences on a cell-by-cell basis against a known flooding event, finding agreement 88-98% of the time. The authors created simple inundation event maps for comparison based upon the two DEM types, although any resolution issues were lost once they re-sampled the LiDAR to thirty-meter resolution. While this study's focus was on comparing DEM resolutions on a cell-by-cell basis, the statistical analysis can be applied to a study of differing resolution floodplains.

2.2 Floodplain Delineation

Floodplain boundaries are delineated based on FEMA guidelines and specifications in the United States (FEMA, 2005). Beginning with the National Flood Insurance Act of 1968 and the Flood Disaster Protection Act of 1973, the US government has played an active role in protecting and insuring flood-prone areas. These two Acts spawned the creation of the National Flood Insurance Program (NFIP) with the subsequent development of Flood Insurance Rate Maps (FIRMs) that are still in use today. With the passage of the National Flood Insurance Reform Act of 1994 and the emergence of digital data and GIS, FEMA's Map Modernization Program was initiated to update floodplain maps. This ongoing project highlights the need for continuously upgrading both floodplain mapping and models as technologies improve.

The existing floodplain boundaries for the entire US were developed using traditional manual cartographic procedures in the 1970s and 80s, prior to the wide application of GIS. Currently, FEMA is upgrading these boundaries through the Map Modernization Program. This program updates the majority of floodplain maps using more advanced methodology and digital data, and it will provide FEMA with the ability to renew and update FIRMs quickly and accurately in the future (Frengs et al., 1999). Flooding incorporated into the FIRMs covers flooding events ranging from 10 to 0.2 percent annual chance for detailed study areas. However, the one percent annual chance flooding event (100 year flood) is still the major focus of the NFIP.

There is a constant need to update and create more accurate floodplain maps. FEMA is in the final stages of a one-billion-dollar Map Modernization Project to update floodplain maps for 92 percent of the US population. Even with these updates, there is still a need to create more accurate floodplain maps as new elevation datasets are constantly being developed. LiDAR is one technique that has changed the way DEMs are created (National Academy of Sciences, 2009). Older techniques of developing DEMs rely on the USGS topographic map contour lines mapping Digital Raster Graphics (DRGs), which in some cases were developed decades ago. Gathering the data and creating the most accurate DEMs can be quite an expensive and tedious process. Even then, technologies such as LiDAR that produce higher resolution DEMs require 'cleaning' away buildings and tree cover before producing an elevation model. As all DEMs introduce some form of error, and all have advantages and disadvantages, studying how a DEM affects a delineated floodplain is important from a technical point of view. This study was conducted to provide insights about how floodplains created using the same process and method but different resolution DEMs compare and contrast.

The Map Modernization program is upgrading the quality of the floodplain maps partially due to updated methodologies, base data, and technology. The effort tries to produce a more accurate floodplain, comparable both to known flooding events and older "Q3" data. Q3 data refers to the manually drawn floodplains originally developed by FEMA using USGS contour maps. These were transformed into digital floodplains by scanning the maps, then rectifying

20

and georeferencing the resulting images. When compared with these existing

Q3 floodplains, newly developed floodplains show some inconsistencies both

between one other and against known flooding events. Unfortunately, the

discrepancies are not consistent; while most are unchanged, some land parcels

are now being added to the floodplain while others are removed based upon new

modeling techniques (Figure 8). Overall it appears that while newer models and

methods of delineating floodplains tend to correspond more accurately with

known flooding events, they too have limitations when comparing modeled

floodplains to known flood events, suggesting that further study is needed

(Aycock & Wang, 2004).
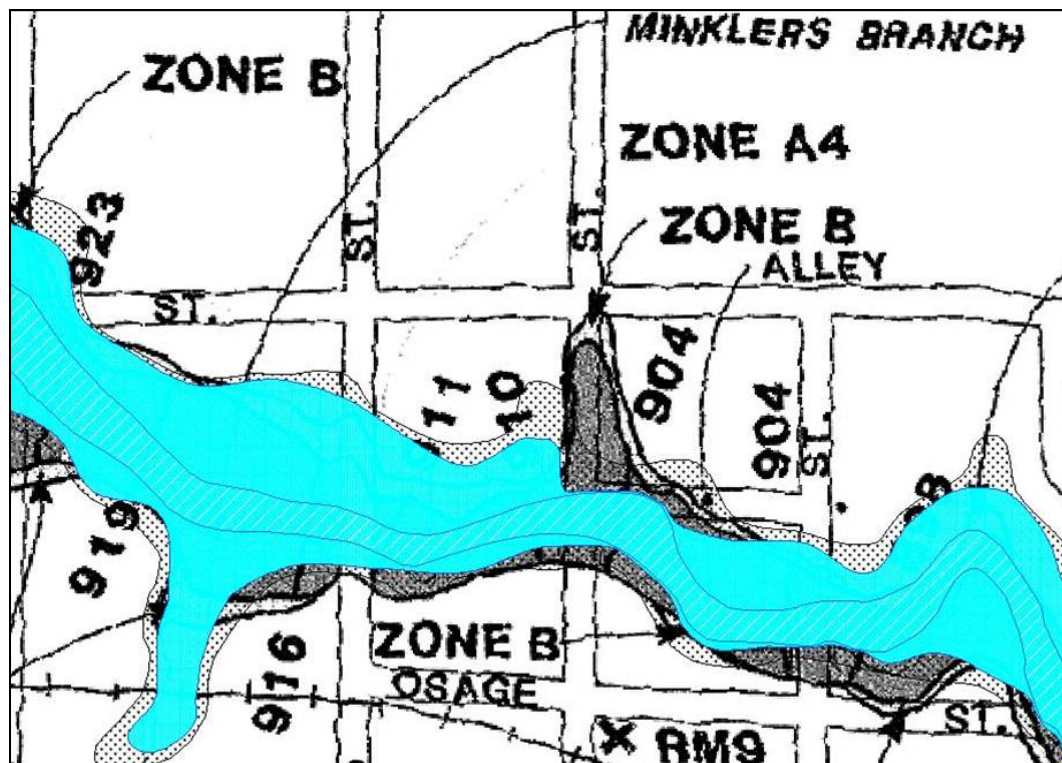
Figure 8: Comparison of 1983 effective and 2008 modeled floodplains
New floodplain overlays the FEMA effective map of Minklers Branch in Plattsburg, MO
Light blue area indicates flooding modeled in 2008 using DEM data
The darker background area represents 100 year flood determined in 1983
The lighter background area represents the 500 year flood event

Aycock and Wang (2004) concluded that even with the best available data and analysis, delineated floodplains were not completely accurate when compared to actual flooding events. This uncertainty makes accurate prediction of flood zones difficult and the methodologies used to create floodplains must be studied and refined (Aycock & Wang, 2004). Because such predictive models are based on numerical modeling and interpolations, further study is required to understand what should be deemed acceptable as compared to known flooding events. Studies of automated floodplain modeling generally acknowledge a need for further study of flooding events and attenuation of models (Pappenberger et al., 2006).

Floodplain delineation usually consists of three sections: hydrologic analysis, hydraulic analysis, and delineation of the floodplain (FEMA, 2005). DEM resolution has an impact on all of these processes. There are also many different modeling techniques for each portion of the process. Different models have strengths and weaknesses, which in turn account for slight variations in results. There is also the added challenge of adapting well-established models such as HEC-RAS to the process. Developed by the US Army Corps of Engineers, HEC-RAS is the dominant hydraulic model used by FEMA, and it has continuously evolved since its initial development in 1964, but it is fundamentally a one dimensional (1D) model. Methodologies allowing for conversion of 1D models into more common GIS formats are becoming more prevalent, intuitive, and accurate (Wheater, 2002; Yang et al., 2006).

It has also been suggested that one of the major uncertainties associated with floodplain mapping is the variety of model choices and combination possibilities (Pappenberger et al., 2006). There are three distinct stages of floodplain delineation, each with a multitude of possible model choices. More studies should be conducted to analyze the variations between model and how extensively these variations affect modeled results. Knowing the extent of these variations allows for the determination of the best modeling technique and for more accurate application of the resulting outputs.

DEM resolution is one variable of the larger floodplain modeling process, but studying the effects resolution has on the modeling process could lead to improved automated modeling. Knowing how the resolution of the DEM affects the output from a model can lead to the development of more accurate input data (or combinations of data) for the entire process.

For the hydrologic modeling, DEM resolution would affect the derived flow path of the stream network and the boundary and area of the watershed delineated. Higher resolution DEMs tend to delineate more accurate stream networks, whereas lower resolution DEMs tend to create shorter flow paths (Horrit and Bates, 2001; McMaster 2002). Alternative methods of DEM creation also introduce some variations, such as the currentness of the data, while the processes that prepare model parameters can alter a model's output. "Filling" a DEM (increasing the elevation of a cell until it is equal to or greater than at least one adjacent cell) alters the surface flow paths, ultimately changing the model's output (Vazquez and Feyen, 2006). One common process during DEM creation

is the process of hydro-enforcing a known stream network into the DEM (Peralvo, 2004). The resulting flow paths more closely resemble reality but are still affected by the DEM resolution and the accuracy of the initial stream network used (Figure 9). Should a stream network be needed, there should be a single ortho-photo rectified network used consistently for all DEM resolutions to limit the introduction of error.
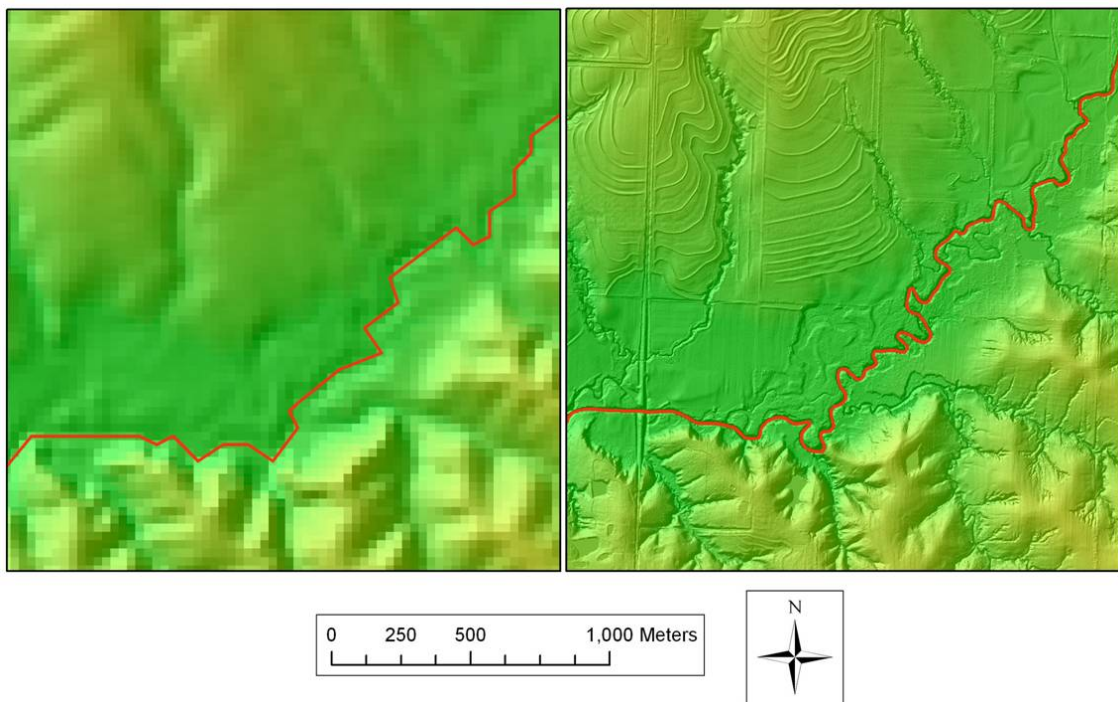
Figure 9: Comparison of thirty-meter and one-meter flow paths
Stream vertexes determined by center of data cells
Smaller cells equate to less disjointed stream pattern

The calculation of water surface elevations is affected to a greater degree because the cross sections identify a much smaller area, due to DEM generalization and cell size, allowing a smaller margin for error. A cross section is a linear representation of the series of elevation points that create the stream

channel profile at a location where a flood water-surface elevation is calculated. In traditional flood studies these points were placed through field verification and analysis; currently, the automated modeling of cross sections derives elevation data directly from the DEM, which means that the resolution of the DEM directly impacts this calculation (Walker and Willgoose, 1999; Hardy et al., 1999).

Cross sections are also linear features, perpendicular to the direction of flow in a stream channel. Due to interpolation of elevations between cells, the different resolution DEMs provide different elevations for the points as well as locational discrepancies (Figure 10). Depending on the hydraulic model used, this variation can directly affect the calculated elevation of water surfaces. Using multiple models such as the HEC-RAS model and the Center for Applied Research and Environmental Systems (CARES) model, the latter of which is based on the Nebraska N-FECT hydraulic tools, it is possible to discern if variations between models are due to DEM resolution or model calculations.
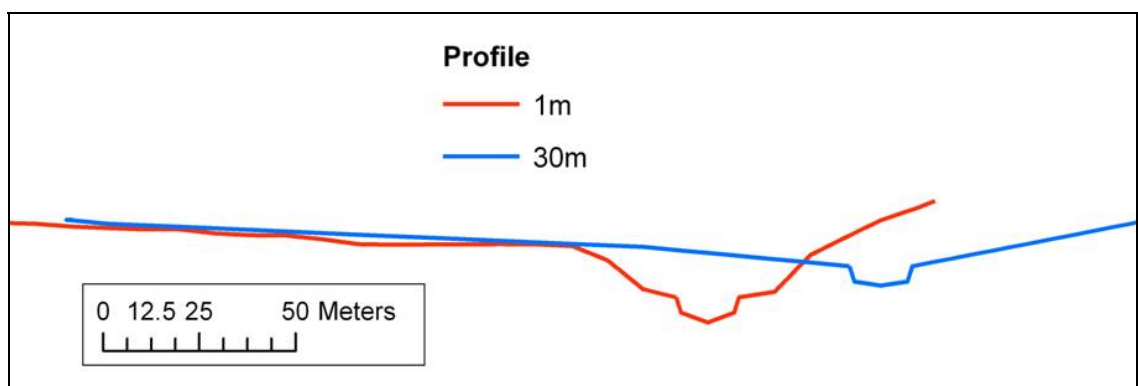
Figure 10: One-meter and thirty-meter cross sectional profiles
These show variations across DEM along same cross section location
Resolution affects both spacing of elevation points and geographic location

The floodplain delineation process takes the known elevations at the cross sections and interpolates a floodplain boundary between them. This boundary can be created using different methods. One of the more common methods used with HEC-RAS models uses a Triangular Irregular Network (TIN) rather than a DEM for boundary delineation. A TIN creates a series of triangles, typically one that satisfies the Delauney criteria, to maximize the interior angles of the triangles and thereby represent the flooded area between the known points (Noman et al., 2003).

When studying the effects of DEM resolution on floodplain modeling it is best to use a floodplain delineation method that relies on interpolations from the DEM itself because using a TIN can completely negate any study of DEM variations in the final floodplain (Noman et al., 2003). Analyzing the effects of resolution on the final product of floodplain delineation rather than the individual portions requires consistent control of possible variables. Studying all possible sources of uncertainty might not be possible, yet studies that consistently account for a single stage's variables lead to better overall understanding (Noman et al., 2003).

Currently it is difficult to determine how delineated floodplains compare with actual flooding events. Different methodologies can be applied to automated floodplain delineation outputs to help discern degrees of accuracy. Two methodologies that can be applied are the linear feature comparison tool (Li et al., 2008) and the overlay fit comparison method (Bates and DeRoo, 2000). The linear feature comparison tool analyzes how line features compare to each

26

other in terms of maximum, minimum, and mean differences. This type of analysis is useful for the comparison of delineated flow paths, watershed boundaries and floodplain boundaries. Another type of comparative analysis utilizes the flooded polygons themselves to study the relationships of geographic overlaps. Both methods provide quantitative options to judge differences between known flooding events and delineated floodplains, or alternatively as a comparison tool of floodplains delineated using different DEMs.

During the course of its relatively short history, automated floodplain delineation has been an imperfect process that needs to be studied further. Many studies cited here have been conducted on various portions of the process, including some analysis on how DEM resolution affects portions of the process. There are also some studies on newer types of elevation data such as LiDAR and how this high resolution affects portions of the modeling process. While these studies have added to the process of understanding how modeling works, it is also important to isolate resolution, among other variables, for the entire automated floodplain modeling process. This thesis attempts to break down the process further and understand the effects of varying DEM resolution on automated floodplain modeling itself.

Chapter 3

Study Area and Datasets

3.1 Study Area

The study area for this study is the Camp Creek Watershed located in Saline County Missouri (Figure 11) approximately seven kilometers east of Marshall (Pop. 12,344). The entirety of the Camp Creek watershed is approximately 69 km^2 while the study portion of the watershed is the most upstream 46 km^2. The study portion of Camp Creek consists of approximately 19 km of stream flowing from the headwaters north of Highway P southward to approximately 1.6 km from its junction with the Salt Fork River. The study area is approximately 50 km upstream from the Missouri River. There are three state highways that traverse the study area, although State Highway P crosses the far northern portion of the watershed, well upstream from the modeled floodplains but does impact calculations for all study drainage basins.

The watershed consists of silt loam soils over a parent material divided equally between limestone and shale. This easily worked fertile soil provides a base for the mainly agricultural usage of the study area. According to the USGS Land Cover Institute's National Land Cover Dataset (NLCD), the watershed is dominated by agricultural cropland (63%) with other major land-cover types in the form of pasture (15.2%) and deciduous forest (15.1%) (NLCD, 2001). One reason this watershed was chosen is its limited impervious surface cover (4.2%). The impervious surface percentage is a combination of both impervious open

use and impervious surface, low intensity. The combined impervious surface is limited to a few minor roadways and scattered dwellings. Therefore the impervious surface should have minimal effect on the floodplain delineation process (Figure 12).
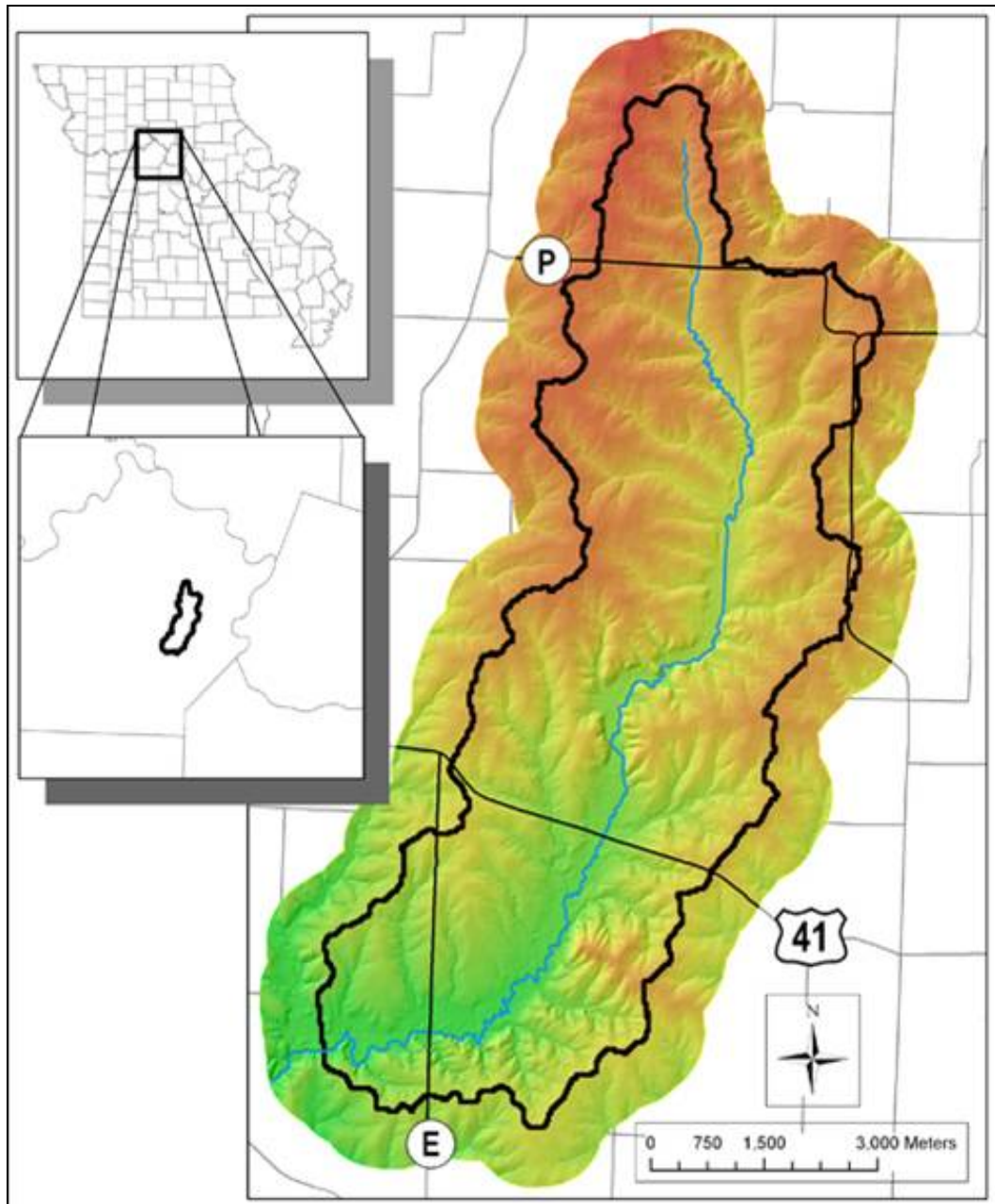
Figure 11: Study Area - Camp Creek Watershed
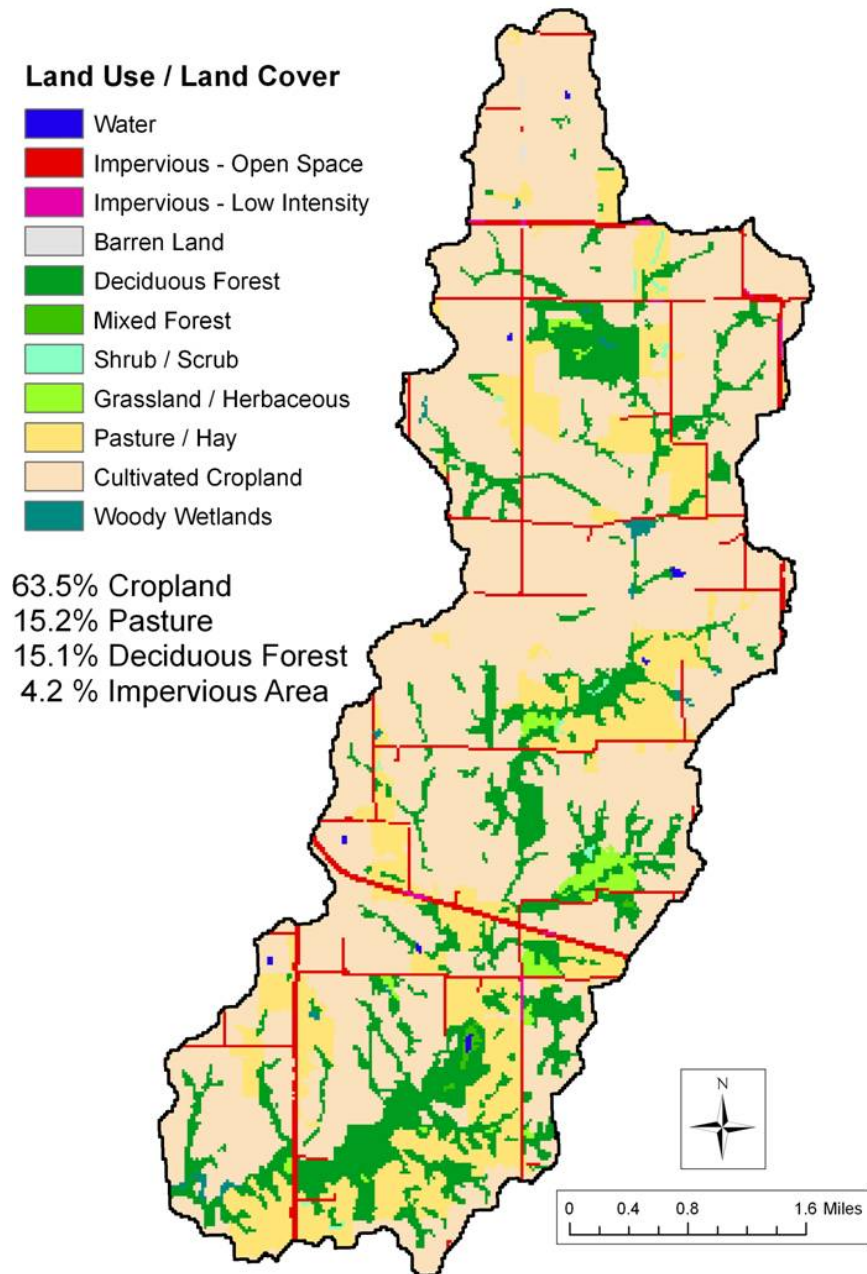
Figure 12: Land Cover - Camp Creek Watershed
National Land Cover Dataset (2001)

The watershed's overall slope varies by DEM resolution. Due to cell size and interpolation of data, changing the resolution of a DEM (or other raster dataset) significantly affects the basic calculations of any derived features (Deng et al., 2007; Shivakoti et al., 2008). Figure 13 shows how changing the

resolution effects the calculation of degrees slope for LiDAR based DEMs.
Higher resolution DEMs tend to show a higher slope average overall due to
smaller cell sizes.
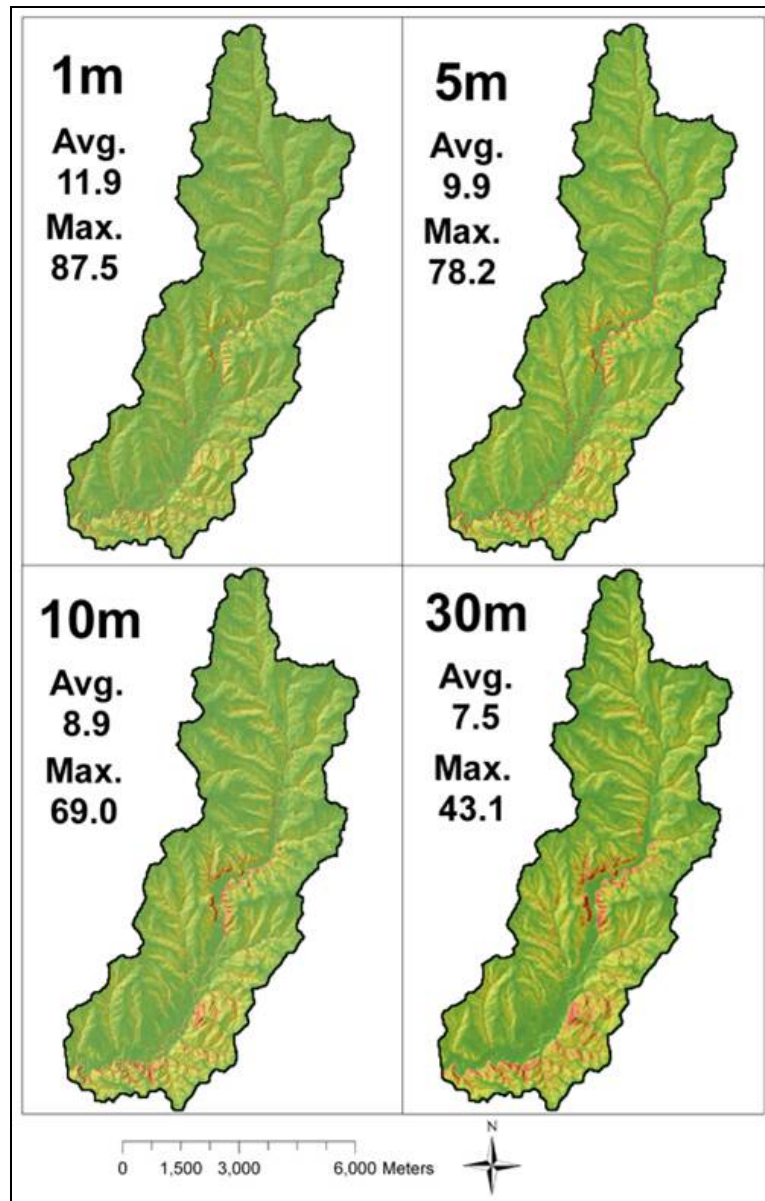
Figure 13: LiDAR slope calculations (Degree slope)

Larger cell size (lower resolution) tends to generalize land surface due to
the loss of smaller steeply sloped surfaces. The one-meter LiDAR slope

calculated for this watershed averaged 11.9 degrees with a maximum slope of 87.5 degrees. By comparison, the thirty-meter LiDAR DEM produced an average slope of 7.5 degrees with a maximum slope of 43 (Figure 13).

Slope calculations based on the USGS DEM followed a similar trend, ranging from an average of 9.6 degrees (maximum of 65) on the five-meter DEM to an average slope of 2.4 degrees (maximum of 16) on the thirty-meter DEM (Figure 14). The comparison between the LiDAR and USGS DEMs is similar for USGS DEMs produced in 2005 at 10 and 5 meters. Additionally, as seen in Figures 13 and 14, a significant discrepancy exists between slopes derived from differing source data thirty-meter DEMs indicating a possible source data inconsistency.
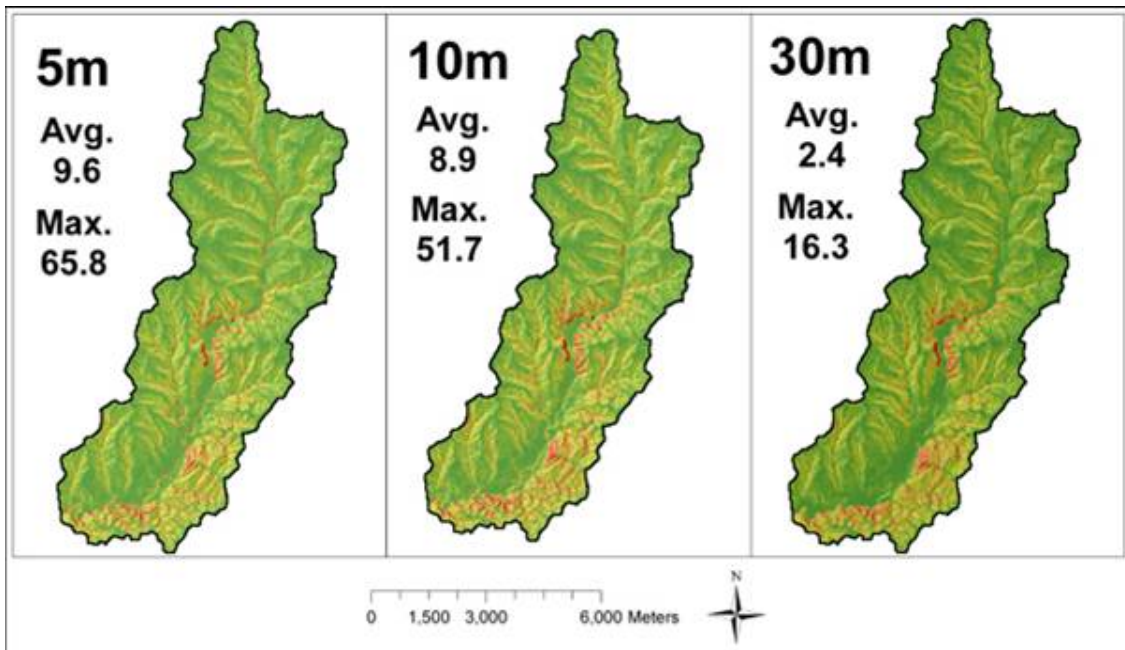
Figure 14: USGS slope calculations (Degree slope)

The study area's low slope, rolling landscape, and low impervious surface area serve as controls for a variety of possible future studies. This study area allows for replication within any non-coastal watershed where a variety of possible variables can be analyzed. This can range from analyzing different models and types of models, DEM resolutions, land cover types, watershed feature types or size, or other variables within the automated floodplain delineation process.

3.2 Datasets

Six different DEM resolutions (1-, 3-, 5-, 10-, 15-, and 30-meters) are chosen to examine the effect of resolution on automated floodplain delineation. DEMs developed from two types of source data are used for this study. One is a set of DEMs derived from one-meter LiDAR data developed in 2007 covering Saline, Lafayette, Carroll and Chariton Counties in Missouri. The other is the existing DEMs available from the USGS developed from 1:24,000 contour data. This includes the statewide 10-meter and 30-meter DEM along with the five-meter Saline County DEM (Table 1).

The LiDAR data was collected and developed into a DEM as a joint project between the United States Army Corps of Engineers (USACE), Kansas City District, the United States Department of Agriculture, and the Natural Resource Conservation Service (USDA-NRCS). The area's proximity to the Missouri River along with the highly active farming community in the area prompted the collaborative dataset development. The agencies supplied the

funding for the data collection and for the development of the DEM by Merrick &

Company using the MARS software for filtering. The dataset was verified using

survey points and has a vertical accuracy to within 0.13 meters.

Table 1: DEM Resolution and Data Source

Type	Source	Resolution	Creation Date
LiDAR	1m LiDAR	1	2007
LiDAR	1m LiDAR	3	2007
LiDAR	1m LiDAR	5	2007
LiDAR	1m LiDAR	10	2007
LiDAR	1m LiDAR	15	2007
LiDAR	1m LiDAR	30	2007
USGS	1:24,000 Contours	5	2005
USGS	1:24,000 Contours	10	2005
USGS	1:24,000 Contours	30	1998

The 30-meter USGS DEM used in the study was retrieved from the

Missouri state data repository (MSDIS) at http://misdisweb.missouri.edu. The

dataset was developed in 1998 for the USGS by the University of Missouri's

Geographic Resource Center using 7.5 minutes digital elevation data provided by

the USGS. The horizontal accuracy is 30 meters with a vertical accuracy RMSE

of 7 meters. The 5- and 10-meter DEMs used for the study were developed by

CARES for the USGS and were completed in 2005. These DEMs were

developed from hypsography digitized from USGS mylar scans of the 1:24,000

scale topography maps. The data indicates a RMSE vertical accuracy of 0.64

meters and a horizontal accuracy for this study area of 3.3 meters. This DEM is

available from the USGS as well as the state data repository (MSDIS). The data

used for this study was obtained directly from CARES with permission of the USGS.

To create the six different LiDAR DEM resolutions, the study area is buffered out 460 meters using ESRI software. This allows extraction of a smaller DEM from the original LiDAR dataset. The one-meter LiDAR DEM is then re-sampled cubically using ArcGIS to create different resolution DEMs. The same buffered watershed area is used to extract the study area DEMs from existing USGS 5-, 10-, and 30-meter data. These DEMs were originally created using 1:24,000 contour data in 1998 and 2005. All data had been developed in NAD 83 UTM Zone 15 projection and a vertical datum of NAVD88.

The USGS DEMs used for this study were previously hydrologically enforced using the National Hydrography Dataset (NHD) (NDEP, 2004). Hydrologic-enforcement is the process where elevation data is forced to recognize a stream network as a low point by incorporating the stream network, along with the contour data when creating the DEM, forcing the DEM to always allow water to flow downhill. This allows the flow of the known stream (NHD) through all portions of the derived DEM, even where the contours would not have depicted any outlets.

For consistency of the study, the LiDAR DEMs are also hydro-enforced. The methodology used by the USGS for hydro-enforcement is replicated using the break lines from the LiDAR as the stream network. The break lines, or interpreted stream banks derived from the LiDAR dataset, are converted into a single channel stream centerline. This stream network is then used to hydro-

enforce each resolution DEM after its conversion from the original one meter LiDAR DEM.

The assumption of this study is that the one-meter LiDAR data provides the most accurate representation of the study area's topography. Because DEMs used for the study were created at different times, and derived from different sources and methods, the study uses the one-meter LiDAR DEM as the reference for all comparisons. How the floodplain modeling environment is affected by the varying data sources is also a focus of the study.

This study used 27 cross section locations for the modeling process. The choice of location for each cross section can introduce human bias into the floodplain process. Cross section locations are determined by technicians and placed according to FEMA guidelines. For the purpose of this study, the effect of any bias is limited by using the same cross sections for all DEM resolution analyses. The FEMA methodology used for Missouri recommends that cross sections be placed approximately every 1.5 meters of vertical elevation change along the stream flow path and at all major constriction points along a stream network. The study area lacked any natural constriction areas but two bridge locations created minor constrictions along the stream network. These placement guidelines mean that additional cross section locations were modeled above the minimum needed for the modeling.

Cross section lines are placed into a coverage format using ArcGIS. The stream network, DEM, USGS contours, LiDAR contours, and Digital Ortho Quarter-Quadrangle aerial photography (DOQQ) provided background data and

imagery. The DOQQs were taken in March of 1997 (published in 2000), allowing

for leaf-off imagery, and they are one-meter resolution. More current imagery

was available for the study area, but the leaf off imagery allows more of the

ground surface identified. Cross sections are drawn as lines between hilltops

along the stream network. They are kept perpendicular to the flow direction of

the stream network and drawn straight where possible. FEMA guidelines allow

for bending of cross sections to follow ridgelines provided the area within the

floodplain remain perpendicular to flow. Cross sections are also drawn along the

main channel with the ends avoiding side channels as necessary.

Cross sections are drawn across the flow path valley, placing the center

point of the channel where the cross section line intersects the stream network.

The USGS has recommended for floodplain analysis in the state of Missouri that

cross section profiles have an artificial channel created for cross sectional

locations that model areas with greater than 5.1 km^2 of drainage. Therefore the

points representing the channel are artificially deepened to create a 1.5 meter

deep by 5 meter wide channel (Figure 15).

This new artificial channel is created by defining points 2.5 meters on

either side of the center point that are identified and labeled as the beginning of

the channel banks. Between the channel center point and the calculated bank

points, an additional set of points is identified to complete the five point artificial

channel geometry. The artificial channels are not created for cross sections with

drainage area smaller than 5.1 km^2. This prevents areas with smaller drainage

basins from having their entire flood being calculated within an artificial channel,

allowing calculated flood surfaces to overtop the natural channel (as defined by

the DEM) and flood surrounding areas.
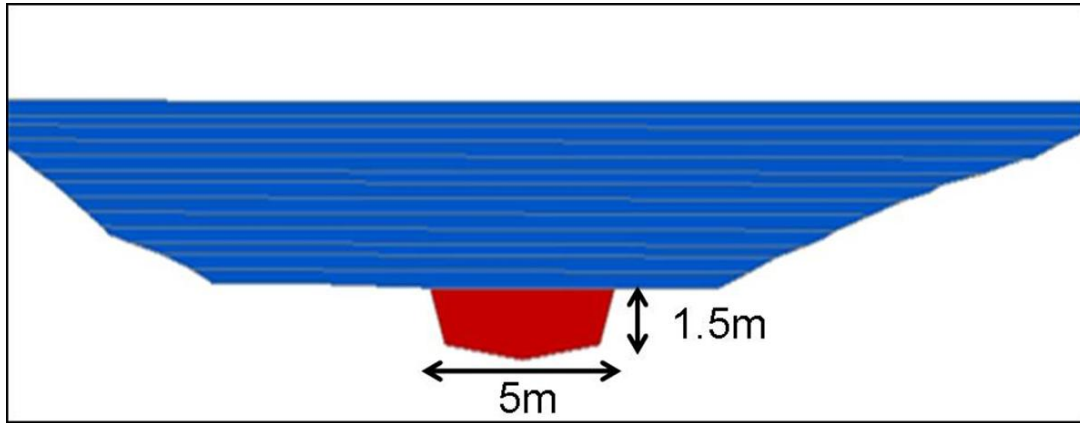
Figure 15: USGS recommended cross section channel addition for Missouri
Artificial channel added to cross sections with drainage area >5.2km²

Chapter 4

Methodology

Floodplain delineation requires multiple input and operating variables. Some variables are not related to the DEM or its resolution. Variations among the types of models chosen for the floodplain delineation process, different settings and parameters affecting these models, as well as varying data sources all affect the resultant floodplain. To study the effect of DEM resolution on floodplains, it is necessary to keep all variables not related to DEM resolution as consistent as possible. Therefore, the study uses the same models and variables at each portion of the floodplain delineation process.

The study examines how DEM resolutions affect the size and shape of a delineated floodplain boundary. This is achieved by resampling a LiDAR one-meter DEM to different resolutions (3-, 5-, 10-, 15-, and 30-meters). The study also assesses the variation between different DEM sources by comparing LiDAR DEMs and USGS DEMs created in 1998, 2005, and 2008. The comparison of floodplains produced by varying resolution and source data is a primary focus of the study.

A secondary focus of the study is to analyze the cumulative effect of resolution on the modeling process. DEMs are used for all portions of the modeling process (hydrology, hydraulics, and delineation) so this research analyzes how variations in DEM resolution affect both individual processes and the entire modeling process. A secondary set of floodplains is produced using a

standardized set of water surface elevations, removing any cumulative variations

in calculated water surfaces, producing floodplains at various resolutions from

the same water surface. Comparing the floodplains produced during the primary

analysis to those produced during the secondary analysis indicates the amount

of cumulative variation injected during the hydrologic and hydraulic modeling

processes.

Additionally, during the study two types of hydraulic models are used to

calculate water surface elevations. The CARES model calculates both

hydrologic and hydraulic results during a single model run. The study produced

two sets of water surface elevations, one from CARES and the other from HEC-

RAS. As HEC-RAS is the more widely used model, only the elevations

calculated using the HEC-RAS model are used to produce floodplains for the

study. As data is produced, a comparison of the calculations from a backwater

model (HEC-RAS) and a normal depth model (CARES) is made. This

comparison is also a tertiary focus of the study.

4.1 Automated Floodplain Delineation

Traditionally, floodplain modeling relied on extensive field work.

Engineers would travel to the study area and extract elevation data across

stream networks, allowing mathematical calculations to be performed.

Historically, these elevations were used in conjunction with topographic maps;

engineers would hand-delineate the floodplains based on engineering judgment

and field observations. The development of digital elevation models allows for

both faster and more accurate floodplain interpolation between observed locations (Noman et al., 2001). High resolution LiDAR elevation models are even taking the place of some field observations when vertical accuracy and cell resolution are within acceptable limits (Simon & Stack, 2009).

The floodplain delineation process for this study includes all three distinct portions of modeling; hydrologic, hydraulic, and floodplain delineation (Figure 16). Hydrologic modeling is the process of determining the runoff and discharge rate of a precipitation event for known locations. Hydraulics refers to the calculation of the water depth and the surface elevation of runoff events at known locations. Floodplain delineation refers to the process used to derive lines along surface features representing modeled floodplain extents. As seen in Figure 16, the study works both within the Geographic Information Systems environment and in the modeling realm of HEC-RAS.

The hydrologic model used for this study also requires an additional process of delineating a stream network (flow path) as an input. The models chosen for this study are consistent with the modeling process used by the State of Missouri to create floodplains for the FEMA National Flood Insurance Program (NFIP). All models used in this study are approved by FEMA during the current floodplain Map Modernization Program. As Missouri has been updating its floodplains using the ten-meter USGS DEM, the CARES models are modified to accept a variety of different DEM resolutions as inputs.
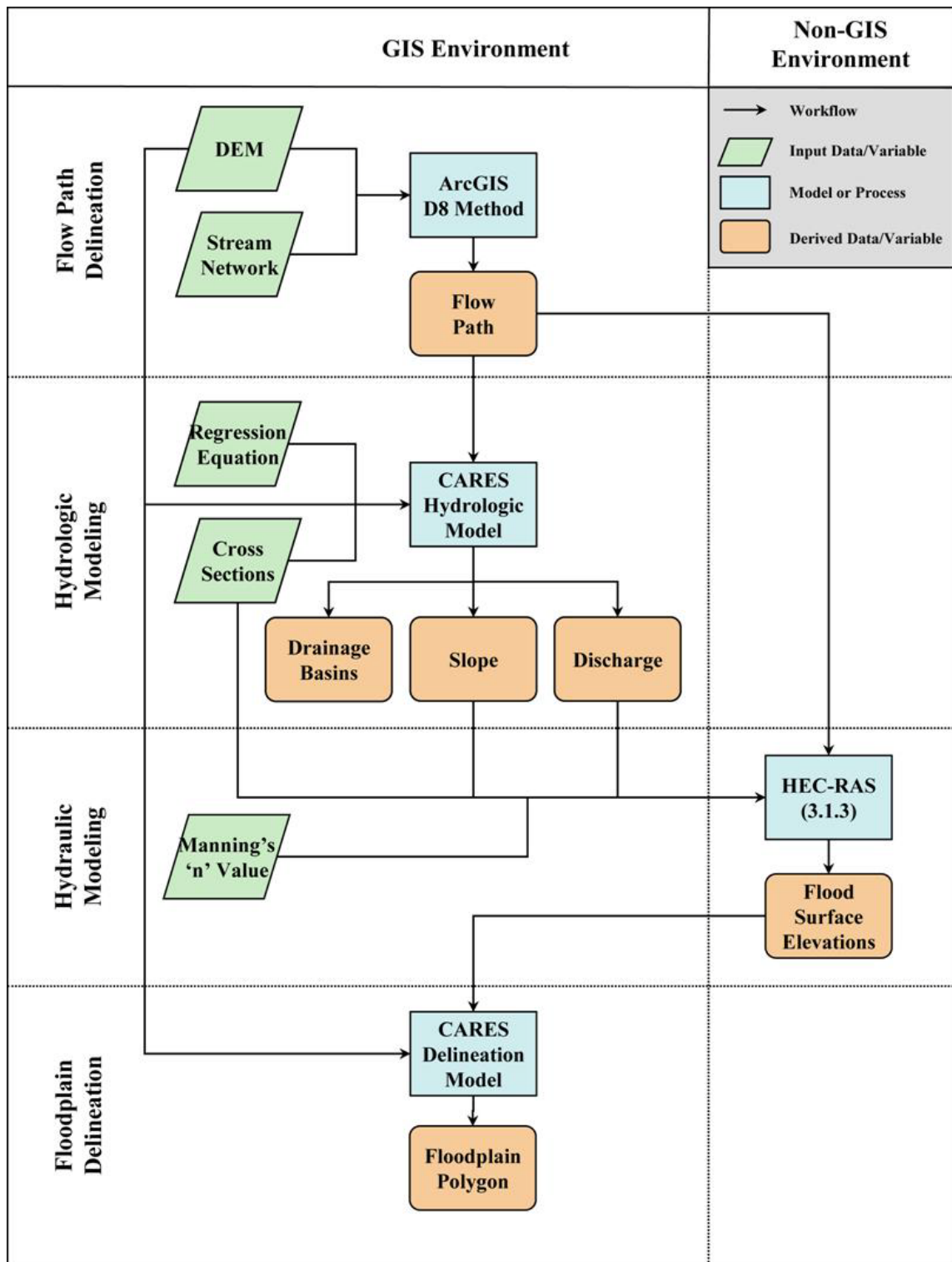
Figure 16: Floodplain delineation processes

Each portion of the modeling process has multiple possibilities for model choice, each with its own strengths and weaknesses. Table 2 lists the nationally approved FEMA models that are acceptable for use with the Map Modernization Project (FEMA, 2005).

Table 2: FEMA nationally approved floodplain models

Hydrologic Models	Hydraulic Models	Delineation
Single Event	1D Steady Flow Models	
HEC-HMS	HEC-RAS	HEC-GeoRAS
TR-20 Win	WSPRO	Floodplain Delineation
WinTR-55	QUICK-2	Toolbox
SWMM 5	HY8	TIN Modeling
MIKE11 UHM	WSPGW	ArcInfo AML DEM
PondPack	StormCAD	Extraction
XP-SWMM	PondPack	
Xpstorm	Culvert Master	
	XP-SWMM	
Continuous Event	1D Unsteady Flow Models	
HSPF 10.10	HEC-RAS	
MIKE11 RR	FEQ / FEQUTL	
PRMS 2.1	ICPR	
	SWMM 5	
	UNET 4.0	
	FLDWAV	
	MIKE11 HD	
	FLO-2D	
	XP-SWMM	
	Xpstorm	
Continuous Event	2D Steady/Unsteady Models	
	TABS RMA2/4	
	FESWMS 2DH	
	FLO-2D	
	MIKE Flood HD	

Each model choice introduces some error and bias, though this is inescapable as all of the model variations and specific models inherently have

strengths and weaknesses. Whether or not there is an optimal method for floodplain delineation is beyond the scope of this study. Further study would be needed to analyze all of the possible combinations of models.

4.1.1 Hydrology

There are two main types of hydrological models: regional regression equations and rainfall-runoff numerical models. Regression equation models use a rainfall-runoff equation for an area to determine the discharge at a given location that constitutes the outlet for a sub-basin of interest. As used in this study the locations are defined by cross sections. This type of calculation uses a regression equation to determine discharge (Q) and is typically followed by a number indicating the return interval for the flooding event being modeled. This study uses the Q100, or calculated discharge for a 100-year flood. The storm surge hydrograph can also be determined by distributed rainfall runoff numerical models for a specific watershed or sub-basin. That type of model uses more specific event inputs such as the data collected from rain gages and soil saturation (Madsen, 2000).

There are multiple hydrologic model choices such as the US Army Corps of Engineers HEC-HMS models, US Environmental Protection Agency's BASINS model, or the engineering firm Watershed Concept's WISE models. The hydrology model for this study is the CARES model. This model was initially an adaptation of the Arc Macro Language (AML) process that eventually became the Nebraska N-FECT toolset. The original CARES methodology was initially

reviewed by the engineering firm PBS&J in 2002. The current CARES model

has since incorporated several improvements and was subsequently reviewed by

the engineering firms of Watershed Concepts and AMEC. While this model

represents hydrology and hydraulics simultaneously, only the hydrology results

are used for this study. This model was chosen primarily because it is designed

specifically for Missouri and calculates a discharge and drainage area at every

cross section location without having to specifically place separate nodes.

Having a calculated discharge at every cross section location enhances the

hydraulic model, allowing for computation of results with less interpolation.

The hydrology portion of the CARES model is based upon regression

equations for the state of Missouri. The regression equations are taken from the

USGS Fact Sheet 015-01 (USGS, 2001). There are three regression equations

used in the State of Missouri. The study area is located within Missouri region

one so the regression equation used is $Q_{100} = 376 \ A^{0.652} \ S^{0.346}$ where Q_{100} is

discharge producing the 100-year flooding event (1% annual chance event), A is

the drainage area in square miles, and S is the watershed slope in feet per mile.

Watershed slope is calculated at points 10 and 85 percent of the distance along

the main channel, divided by the distance between the points.

The CARES model requires a basic stream flow path to be derived from

the DEM. This is accomplished using the standard D8 method in ESRI Arc

products using the ArcInfo toolset through a series of commands in AML. Even

though the DEMs are previously hydro-enforced, the standard D8 method

requires removal of all low lying areas with no outflow (sinks). The DEM is

prepared by accounting for all sinks, filling them to the lowest area and surrounding them so that all portions of the DEM have a single outflow location. The DEM is then converted into a flow direction grid that tracks how water would ideally leave the cell and enter the next 'lower' cell along the surface. A grid cell has only eight possible outflow directions (i.e., neighboring cells) and the flow direction process assigns the value to the cell based on its outflow direction. This networking of cells creates a flow accumulation grid, assigning values to cells by tallying the total number of upstream cells. This number represents a number of cells but can also be converted into drainage area by calculating the number of cells and multiplying that by the DEM resolution.

Using the D8 method, a flow path network is derived from the DEM that shows the idealized flow path for water flowing along the surface of the DEM (Tribe, 1992). For this study the flow path for Camp Creek is extracted from the DEM for all areas greater than 0.25 square kilometers of drainage. Even though floodplains are only delineated for areas with greater than 2.6 square kilometers of drainage, a longer flow path helps calculate variables such as cross section drainage area and slope more accurately. This process is repeated for each DEM resolution, resulting in nine distinct flow paths (6 LiDAR and 3 USGS). While each flow path served as the base for floodplain delineation of individual DEM resolution, the one-meter LiDAR stream network served as the reference when comparative analysis is conducted (Figure 17).
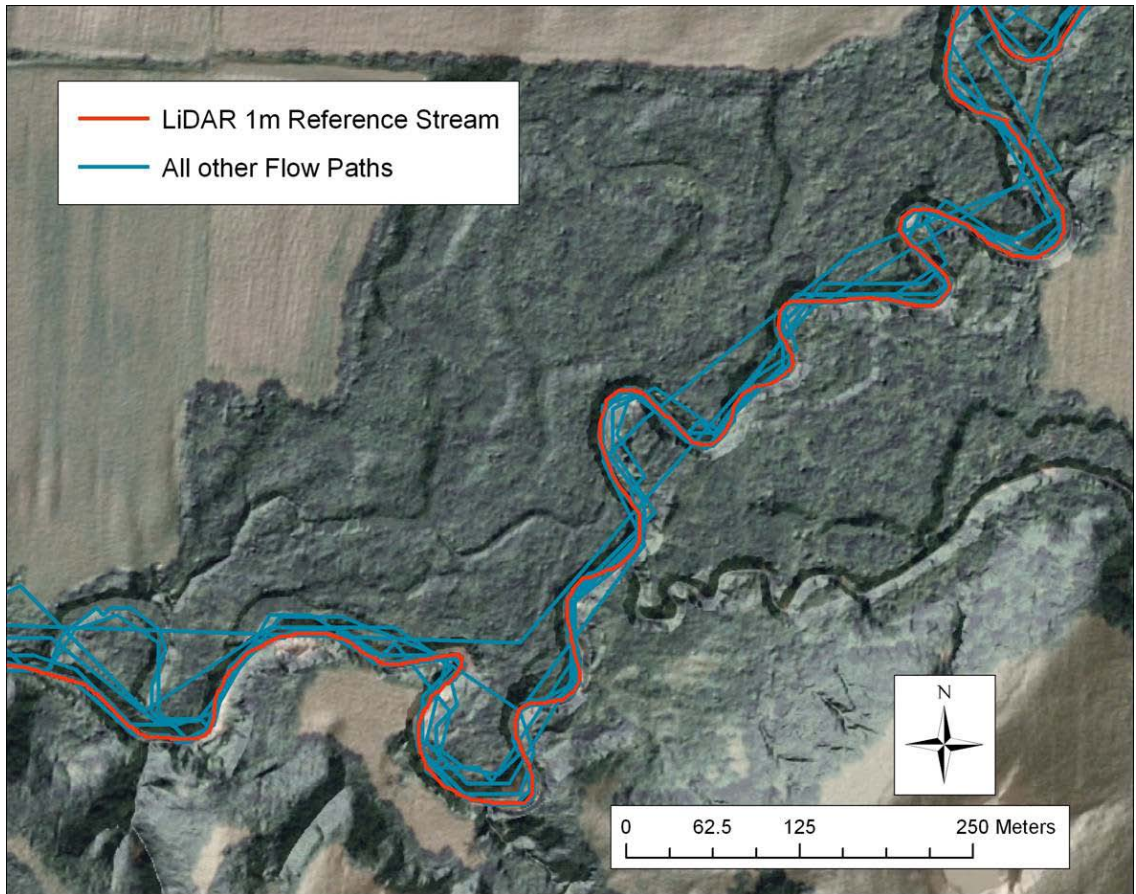
Figure 17: Locations of the nine study flow paths

One advantage of using the CARES model is that it does both hydrologic and hydraulic analysis simultaneously. Incorporating the hydraulic features (cross sections) into the hydrologic modeling allows all hydrologic calculations to be performed at each cross section location. Watershed area, drainage area slope, local slope, and discharge are calculated at each of these cross section locations, providing more accurate input data for the hydraulic model. Having a known discharge, cross sectional profile, drainage area, slope, and channel location at all calculation locations resulting in more accurate flood surface elevation calculations.

To calculate the discharge for each cross section, a watershed drainage area for each location is derived. The CARES model uses all cells that touch the cross-section to determine the drainage area. The output representing each location's drainage basin is a grid that is converted to a polygon. The delineation of the watersheds allows for the comparison of different DEM resolutions on watershed delineation (Figure 18).

Figure 18: Twenty-seven cross section locations
Black lines represent the cross sections used for the study
Red lines represent drainage basins derived for each cross section
Focus area shows example of 'bent' cross section.

As seen in Figure 18 there are some cross sections that are not always placed perfectly straight as is preferable for hydraulic modeling. For this study the cross sections are set to be perpendicular while within the channel valley but are allowed to bend once they leave the floodpl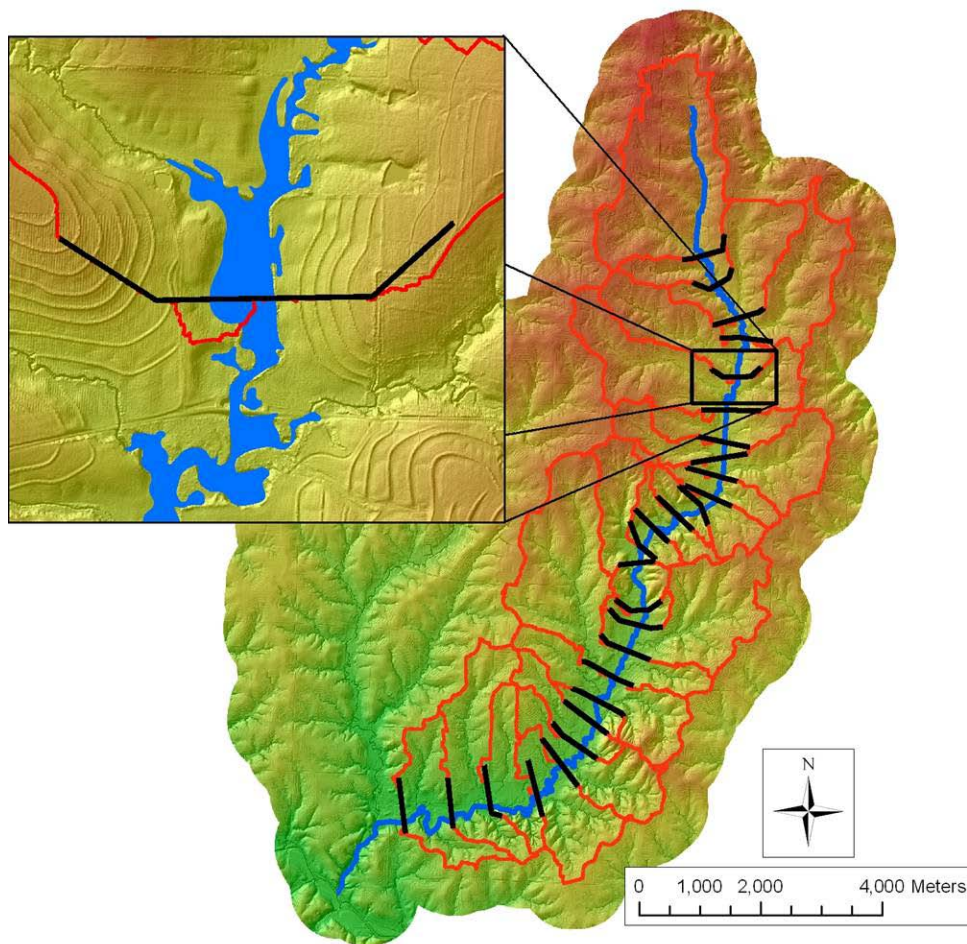ain area. This is done to ensure coverage for flooding calculations at all various resolutions and helps set the watershed boundaries along the ridgelines.

Two slopes are calculated during the hydrology portion: the "85-10 stream slope" and the cross section "local slope". The 85-10 slope is calculated for each cross section by locating two locations along the flow path. The first point is 10 percent of the total upstream distance of the flow path from each analysis point to the watershed boundary and the second is 85 percent of the total upstream distance. As the flow path is not derived to the watershed ridge line, this portion of the model used the watershed polygon, along with the flow path to determine the furthest point upstream from the cross section within its watershed. The final 1/4th of a kilometer of drainage flow path is calculated as a straight line distance to the far point of the watershed. This straight-line distance is then added to the distance along the flow path itself to determine the total flow path distance. Taking this total length, the model then queries the elevations of the DEM at both 10 percent and 85 percent locations (Figure 19). Using these two elevations, the slope for each cross section's contributing area in 'feet per mile' is calculated as outlined by USGS Fact Sheet 015-01 (USGS, 2001).

The CARES model also calculates a localized slope for each cross section. This is calculated by following the flow path geometry downstream 304

49

meters from the intersection of the cross section and the stream flow path. By

querying the elevation of the DEM at that location, along with the elevation at the

cross section itself provides the model with the local slope. This local slope is

only used for hydraulic calculations within the CARES model.

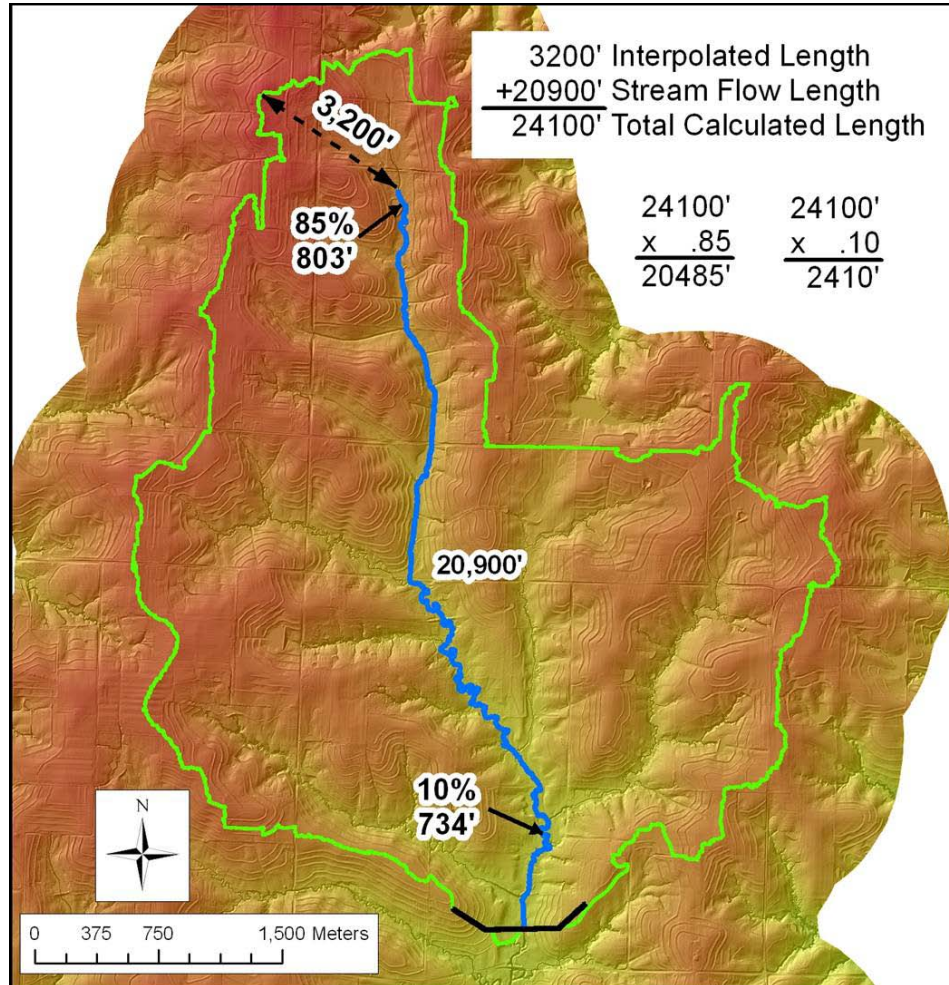**Figure 19: 85/10 Slope calculation methodology ('Slope per Mile')**
Stream length is known from cross section upstream to known point
Straight line distance to highest elevation point is added to length
Elevation extracted from DEM at 85% and 10% total stream length

The ability of the CARES model to produce all results for hydrology

modeling at every cross section location, along with providing a method for

extracting exact elevations along the cross section itself provides a large amount of data that are imported directly into the hydraulics portion of the delineation process without relying on extrapolation. The model also provides a large array of data that helps identify how DEM resolutions affect portions of the hydrologic modeling process.

4.1.2 Hydraulics

The calculation of water heights needs to incorporate many features such as the amount of water or channel area, its rate of conveyance or discharge, the profile of the valley surface or wetted perimeter, stream slope, and the types of land cover at that location. The types of land cover are typically represented by a Manning's 'n' value so adjustments can be made for resistance by materials outside the normal stream channel. The flood height is determined by combining these inputs together to calculate the water surface elevation that would be required to allow passage of the calculated amount of water.

Hydraulic models include backwater, network, and normal depth models; the largest percentage of hydraulic modeling is done using the Army Corps of Engineers HEC-RAS family of models (FEMA, 2005). HEC-RAS is a backwater model. For this thesis backwater model refers to a model that adjusts the calculated normal depth to represent a water surface that is consistent with other upstream or downstream cross sections. The group of models are termed backwater as the majority of water surface calculation adjustments are due to water backing up from a constriction downstream although this is not always the

51

case for adjustments. There are many other model options such as the USGS's

WSPRO model or the EPA's SWMM model. The CARES model, used in this

study solely for hydrology calculations, is an example of a normal depth

hydraulics model. This type of model refers to one where water surface

calculations always equal normal depth calculations; all cross section elevations

are calculated independently, allowing for the possibility that downstream

elevations might be higher than upstream results. The HEC-RAS model, version

3.1.3, was chosen for the study as it is the dominant hydraulics model.

The HEC-RAS model required an import file that is created from the

CARES hydrology model results. HEC-RAS imports all stream vertex locations

and cross section vertexes and elevation profiles as well as discharge at all cross

section locations and slopes of all cross sections. The point elevations along the

cross section profile determine the water heights at that location. A single set of

cross sections is used for all computations, but the points where the elevation

values are chosen to describe the cross section profile are based on the DEM

resolution. The model that constructs cross section profiles has a limitation

preventing it from gathering elevation data from more than a single point within

the same DEM cell. The model is set this way to prevent the cross section profile

from giving false areas of level ground. The points are extracted using the

hypotenuse length of an isosceles triangle ($X\ 2^{0.5}$) where the congruent sides

equal the DEM resolution; hence the one-meter cross section profile has more

cross sectional elevation points than the thirty-meter cross section profile. All

points are gathered using X, Y, and Z values for input into HEC-RAS. Figure 20

shows how the DEM resolution affects the profile shape and the resultant
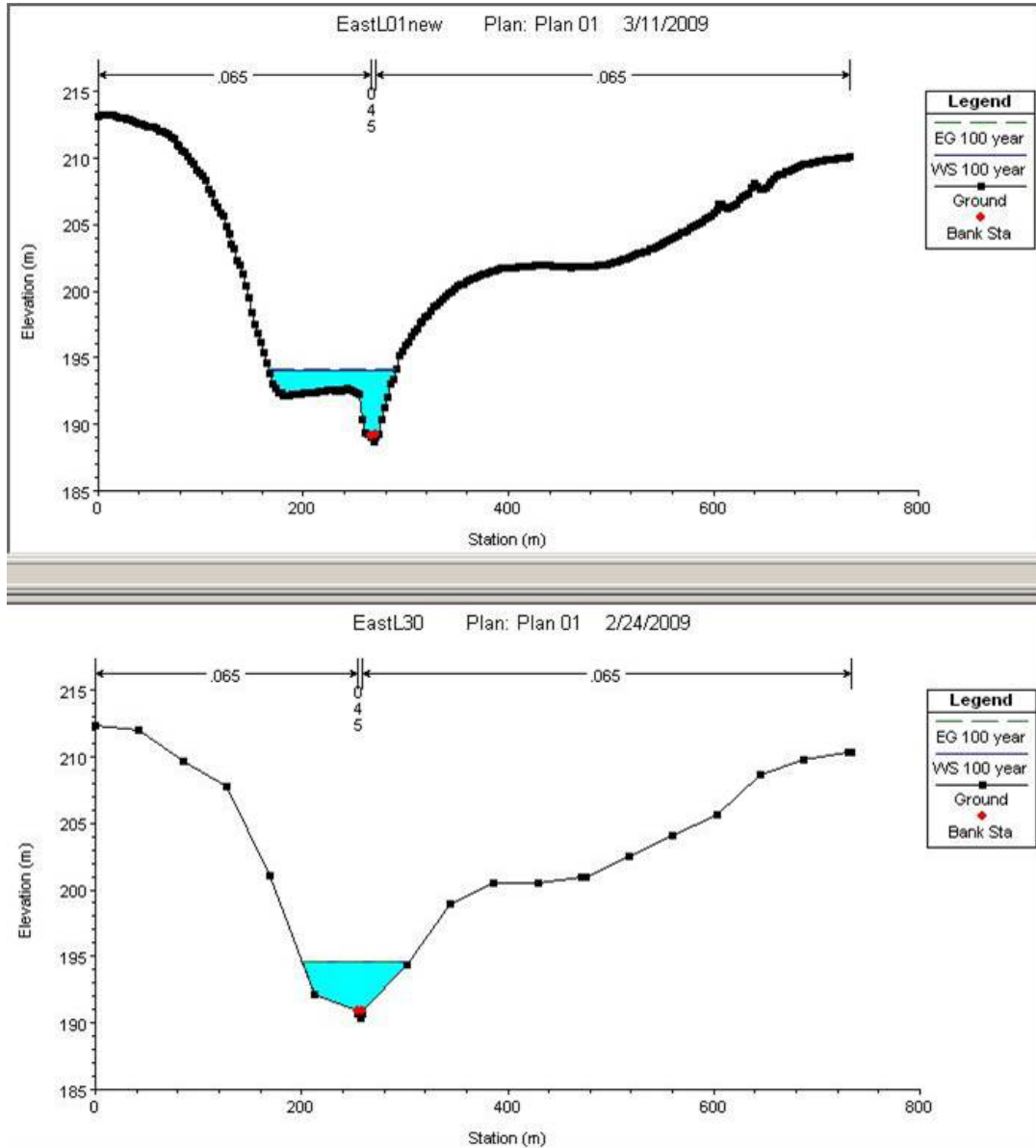
hydraulic calculations.

Figure 20: Comparison of cross section profiles in HEC-RAS
Points are extracted for cross sections based upon DEM resolution
Top: 1m cross sectional view across valley (points every ~2m)
Bottom: 30m cross sectional view across same valley (points every ~45m)

One variable required in HEC-RAS is the Manning's "n" roughness coefficient. This coefficient is a number greater than zero but less than one that represents the roughness along the banks of the stream network. For general approximate mapping a standardized value of 0.065 for banks and 0.040 for channels is used for Missouri region one. While these might not be completely appropriate for the actual roughness for the entirety of the study area, the value is kept constant for all modeling runs. This consistency of input should allow for all model results to be compared and if there is any effect, all outputs will be affected equally.

Once all variables are collected, an input file is prepared for HEC-RAS version 3.1.3. This file provided HEC-RAS with all information needed to calculate the flood water heights. The cross section and stream location information is entered into HEC-RAS using the Geometric Data Editor while the discharge information is entered as existing flow data (USCoE, 2002). Once all data is entered and visually verified, a simple steady flow analysis is run using a subcritical flow regime.

Resultant calculations and profiles are checked for anomalies and then exported from HEC-RAS back into the cross section file. In this subcritical flow calculation checks are done to confirm that no downstream water surface is higher than any upstream calculation. HEC-RAS also produces a stream profile that can be used to check for errors visually. Once all cross section data are calculated and verified, results are then exported from HEC-RAS back to the original GIS cross section file for floodplain delineation. All DEM resolutions are

calculated individually, and all calculations stored in the cross section file for further analysis.

4.1.3 Floodplain Delineation

Floodplain delineation is the process of using flood depths (water heights) and elevations calculated by hydraulic models to delineate floodplain boundaries. The boundary represents the extent of a flood based on the level of flooding modeled. Floodplain delineation is typically done using Triangular Irregular Networks (TIN) to convey output coordinates from the hydraulic model into a floodplain polygon. As this study's primary focus is to analyze the affects of varying DEM resolution on floodplains, it uses another CARES model that creates floodplains by directly using a DEM.

The CARES method relates all hydraulic calculations back to the original cross section lines in ESRI coverage format (Figure 21). With the calculated results imported back to GIS format, they relate more directly to other formats such as the DEM that exists as a raster dataset. The AML used for this process interpolates the distance between multiple points along neighboring cross sections, dividing evenly the distance into 0.15 meter elevation intervals. These interpolated cross sections are referred to as 'dams' as they effectively block interpolated water levels behind them to an elevation equivalent to the 'dam.' For this study, the number of interpolated cross sections is increased to include a 'dam' for every 0.06 meter change in elevation (Figure 21).
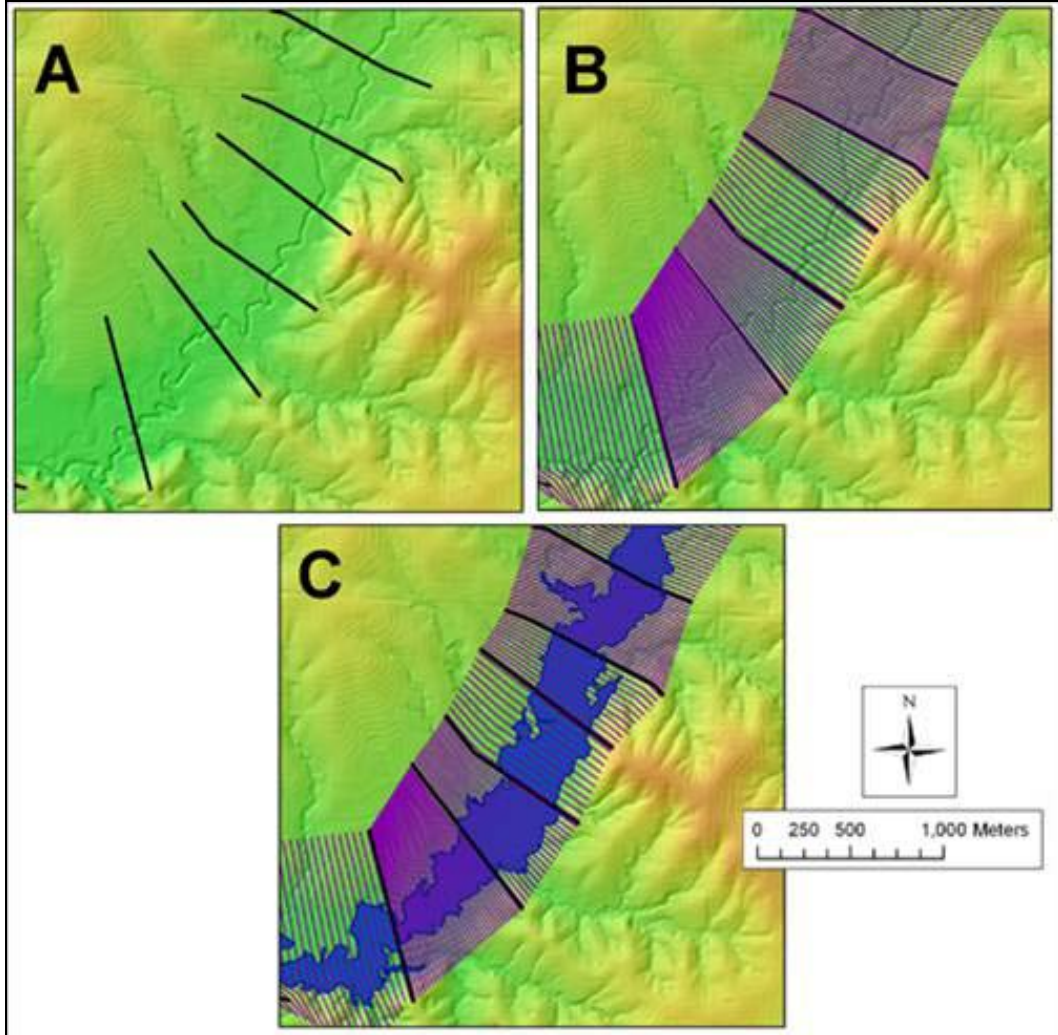
Figure 21 A: First cross sections are placed
B: Cross sections are used to interpolate 'dams'
C: Cross sections and interpolated 'dams' are used to delineate floodplains

After the 'dams' are inserted into the layer, they are converted into raster format as a series of cells consisting of a constant elevation. These are then merged with the existing DEM and the newly formed DEM is then filled so that the 'dams' capture all cells in between as a water surface. After the filling, the new DEM is compared with the original DEM and areas that are changed represent the flooded area. This flooded area is then extracted from raster format and converted to a polygon feature representing the modeled floodplain

(Figure 21). This process relies on the accuracy of the DEM for boundary

detection and can result in various outputs when cell sizes change.

4.2 Comparison of Model Outputs

After all floodplains are delineated for different DEM resolutions, statistical

analysis is performed to compare the model outputs. In order to fully assess the

DEM resolution effects, not only are the floodplain boundaries compared, but

also the different stages of output such as stream flow paths and watershed

boundaries. All study features are GIS grid, linear, or polygon features and all

analyses used the one-meter LiDAR feature as the reference.

Three methods are used for comparison. All linear feature comparisons

use a tool that uses ESRI products to compare features. Raster data sets are

analyzed using an overlap fit analysis, and polygon features are analyzed using a

flooding probability index. The linear feature comparison tool is used to analyze

differing stream flow paths, watershed boundaries and resulting floodplain

boundaries. It provided a way to quantitatively describe the differences between

the nine resolution outputs. As floodplains are the focus of the study, the overlap

fit analysis and the flooding probability index are used to analyze the areas of the

resultant floodplains (Table 3).

Table 3: Study Analysis Methods

Data	Linear Feature Comparison	Overlap Analysis	Probability Index
Stream Network	X		
Drainage Basins	X		
Floodplain Boundary	X		
Floodplain		X	X

4.2.1 Linear Feature Comparison

In many fields of study there often is a need to provide a statistical assessment for the correspondence (or "goodness of fit") between linear features. This could be conducted between model and field data, or in the case of this study, the differences between multiple modeled outputs. An Automated Proximity and Conformity Analysis (APCA) tool was developed as a method to quantify the correspondence of field observations and modeled geological phenomena, and it provides a useful analysis tool for many other fields (Napieralski et al., 2006). Li et al. (2008) provided a revised APCA tool, providing statistical results of linear features in maximum, minimum, mean distance, and standard deviation. The APCA tool uses GIS buffers to compare proximity and parallel conformity between line features and allows for a common level of agreement between features to be analyzed statistically.

The APCA tool is used on the stream data, cross section drainage basin boundaries, and floodplain boundaries (bank data). The line features are compiled into one ESRI shape file for analysis. The stream networks and watershed boundaries are analyzed once while the floodplain boundaries are

analyzed twice. There are cases where the left bank of one trial run is actually

closer to the right bank of the master dataset. Therefore different bank

boundaries are analyzed separately.

4.2.2 Overlap Fit Analysis

The study also presented a need to express coincident portions of polygon

features. To address this analysis an overlap-fit analysis introduced by Bates

and DeRoo (2000) is used when analyzing similarities between two floodplain

features. The overlap analysis is calculated by dividing the intersection of two

areas by the union of the same features and then multiplying that number by 100

to produce a percentage. A Fit Percentage of 100 indicates a situation when the

two areas are completely coincidental (Bates and DeRoo, 2000). The further the

percentage is from 100, the less coincidental the two features are. In the

analysis of the floodplain polygons, the one-meter LiDAR floodplain is always

used as the reference floodplain (M1), while M2 represented the modeled

floodplain being calculated (Figure 22). The formula is:

$$\text{Fit (\%)} = \frac{M1 \cap M2}{M1 \cup M2} \times 100$$

This percentage provided a quantitative comparison based on how each

subsequent floodplain overlapped the reference one-meter LiDAR DEM

floodplain.
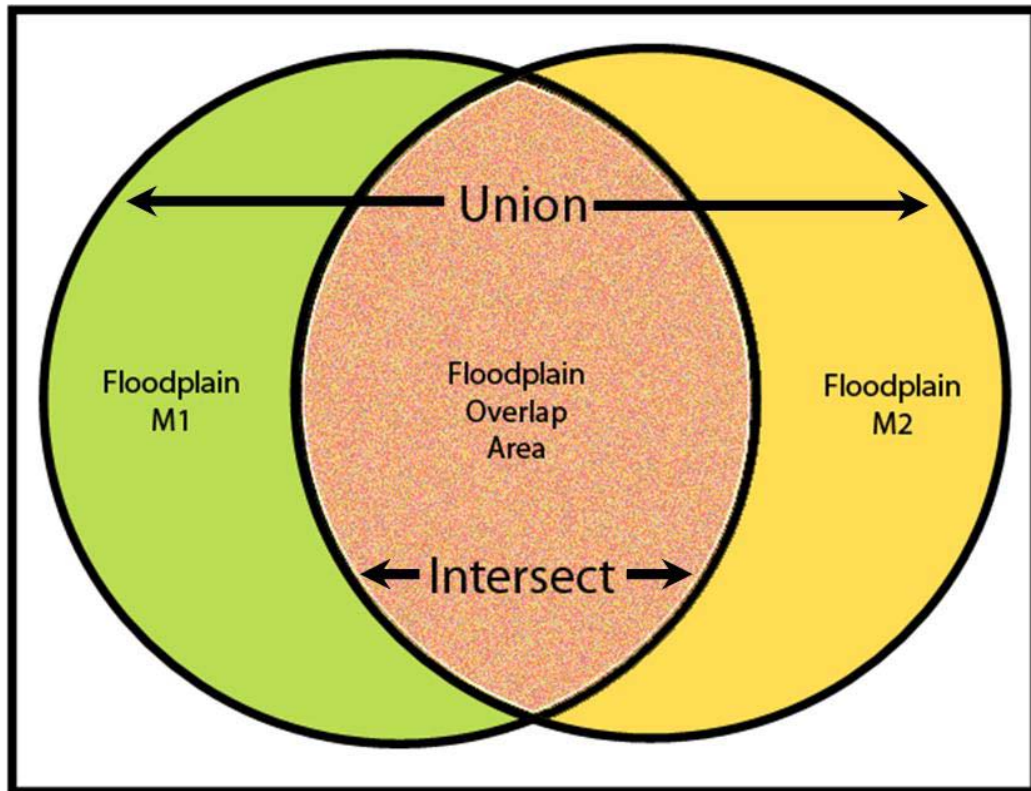
Figure 22: Overlap fit percentage
Intersect / Union * 100

4.2.3 Floodplain Probability Index

One final comparison method is used to analyze the probability of an area belonging to the floodplain. Assuming that different DEM resolutions are independent of each other, and equally important, by overlapping all floodplains from different DEM resolutions, it is possible to determine the probability that an area will be delineated as 'in' the floodplain. If an area (pixel) belongs in the floodplain, regardless of DEM resolution, the probability is 100 percent; if the area always falls outside the floodplain, its probability is zero (Figure 23). The study floodplains are stacked to create a flooding probability index.
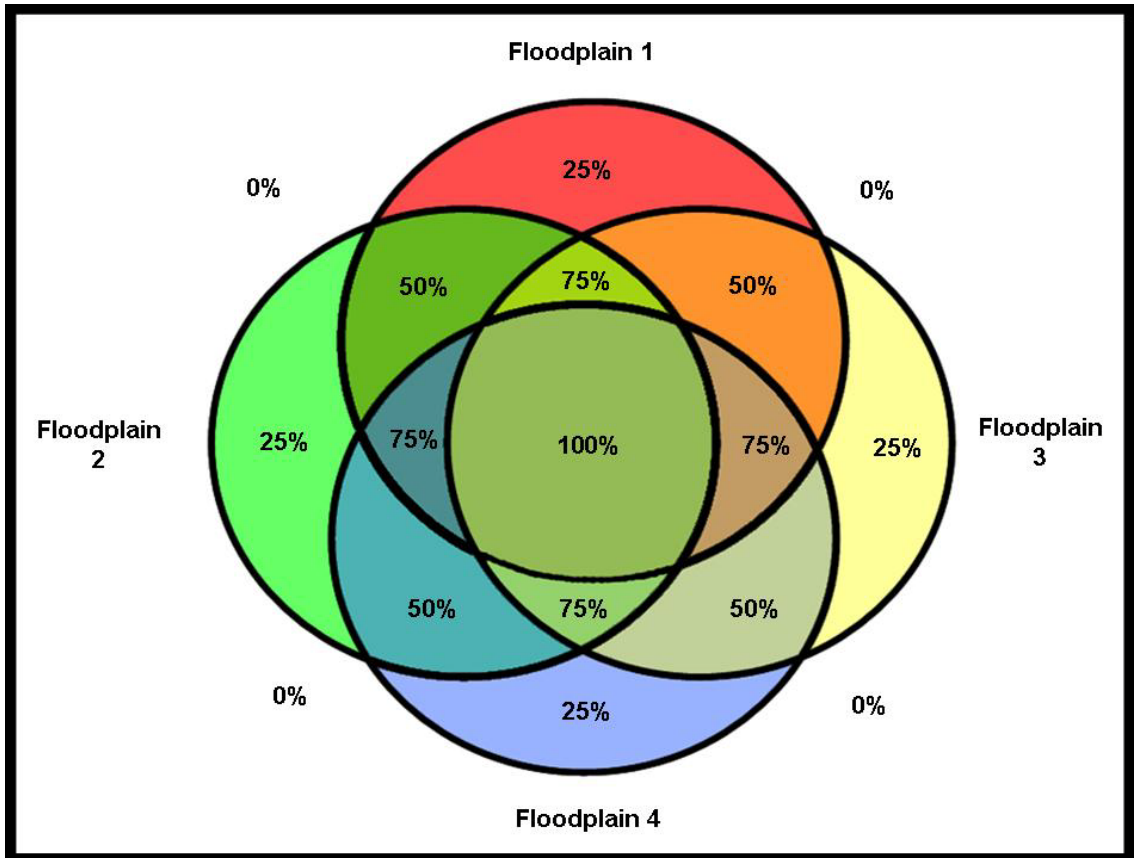
Figure 23: Floodplain probability index

The index indicates the number of floodplains that overlap any given spot within the study area, which provides a overview of how well aligned the flooding events are. The floodplain polygons are converted into one-meter resolution raster images and aligned to sit directly above one another. Null values within these grids are converted to zero while the feature is identified by the number one. The probability index map is then created using a raster calculator, adding all cells together with values ranging from zero (no flooding) to nine (all floodplains overlap). Once the index rating is complete the total area (pixels) that share each rating are combined. Dividing the total area by all cells with a value

greater than zero indicated the percentage of area flooded for that floodplain rating.

4.3 Cumulative Model Effect Analysis

A secondary focus of the study is to analyze the cumulative effects of DEM resolution on the entire modeling process. Inconsistencies introduced during the stream delineation, hydrology, and hydraulic modeling are carried forward through the floodplain delineation process. This circumstance results in a cumulative effect of the DEM resolution on floodplain delineation process: each process introduces slightly more variation based upon varied inputs. A second set of floodplains is created using the resultant floodplain calculations for the reference one-meter LiDAR dataset. This secondary set of floodplains all used the common elevation calculated during the primary modeling for the one-meter LiDAR floodplain. These secondary floodplains allowed for a comparison of floodplains using resolution variation at all model steps against floodplains that had variations introduced only during the delineation process (Table 4).

Table 4: Primary vs. secondary analysis

Study	Flow Path Delineation	Hydrologic Model	Hydraulic Model	Floodplain Delineation
Primary Analysis[a]	X	X	X	X
Secondary Analysis[b]				X

[a]DEM resolution introduces variation at all stages
[b]By using common water surface elevations, only the effect of DEM resolution on the floodplain delineation is examined.

By standardizing the elevation input into the floodplain delineation model, an analysis of how much cumulative variation had been introduced during the previous stages is possible. The secondary analysis used the same 'dams' for every model run so DEM resolution is the only variable introduced. The floodplains still varied but this secondary analysis allowed the amount of variation introduced by this model to be assessed. The secondary floodplains followed all the same analysis processes and are analyzed using the APCA tool on the floodplain boundaries, the overlap fit percentage, and flooding probability index. The results of the secondary analyses are then compared against the primary results to identify the cumulative modeling effects of DEM resolution variation.

4.4 Data Source Variations

This study also analyzed the amount of variation introduced between DEMs created from two different data sources. The source data for the LiDAR and USGS DEMs are created from different sources, processes, and at different times. This introduced the possibility of location and temporal differentiation as a possible error. All analysis results are examined for the six LiDAR DEM resolutions both individually from, and collectively with the USGS results. Analyzing the six LiDAR DEM resolution outputs allows for a stricter comparison of resolution, while using the three USGS DEMs allowed study of how DEM availability has evolved and affected floodplain delineation. Studying the six LiDAR datasets results against only each other eliminates variation due to source material. Using all the same analysis tools, the comparison of the 5-, 10-, and

30-meter LiDAR resolutions results with the USGS counterparts identify the

amount of discrepancy introduced by varying source data.

Chapter 5

Results

5.1 Stream Networks

All six LiDAR DEMs (1 to 30-meter) were derived from the same dataset (1-meter LiDAR DEM) and then individually hydro-enforced using the same stream network. The raster resolution had a profound effect on the calculated stream flow paths. Higher resolution DEMs provided a flow path that more closely resembled the actual stream network. All stream flow paths had similar upstream and downstream endpoints but there was slight geographic variation as flow path endpoints were dependent upon cell locations for each DEM. To account for this, all flow paths were clipped using a line perpendicular to the stream valley nearest the location where all nine flow paths existed at both the upstream source and downstream pour points. While these locations were all similar, the length of the streams varied greatly (Table 5). Variation of flow path location and length indicated that the type of discrepancies will have a cumulative effect on the entire modeling process. Flow path lengths are used as variables during calculations beginning with the hydrologic model's slope calculations.

The effect of DEM resolution is not only reflected as the length difference but also the horizontal offset of the flow path (Figure 24). Using the ESRI D8 method, flow paths were derived by drawing straight lines between the centers of each cell on the flow accumulation raster. Higher resolution DEMs have more

cells, allowing for more vertex points. The much higher ratio of vertexes allows

the modeled flow path to resemble the actual stream network more closely.

Table 5: Stream flow path lengths		
DEM	Length (km)	Offset Percentage*
LiDAR 1m	18.33	0.0
LiDAR 3m	18.12	-1.1
LiDAR 5m	17.83	-2.7
LiDAR 10m	17.30	-5.6
LiDAR 15m	16.51	-9.9
LiDAR 30m	15.24	-17.9
USGS 5m	17.90	-2.4
USGS 10m	17.85	-2.6
USGS 30m	14.52	-20.8

* (Stream length - Reference stream length) / Reference stream length
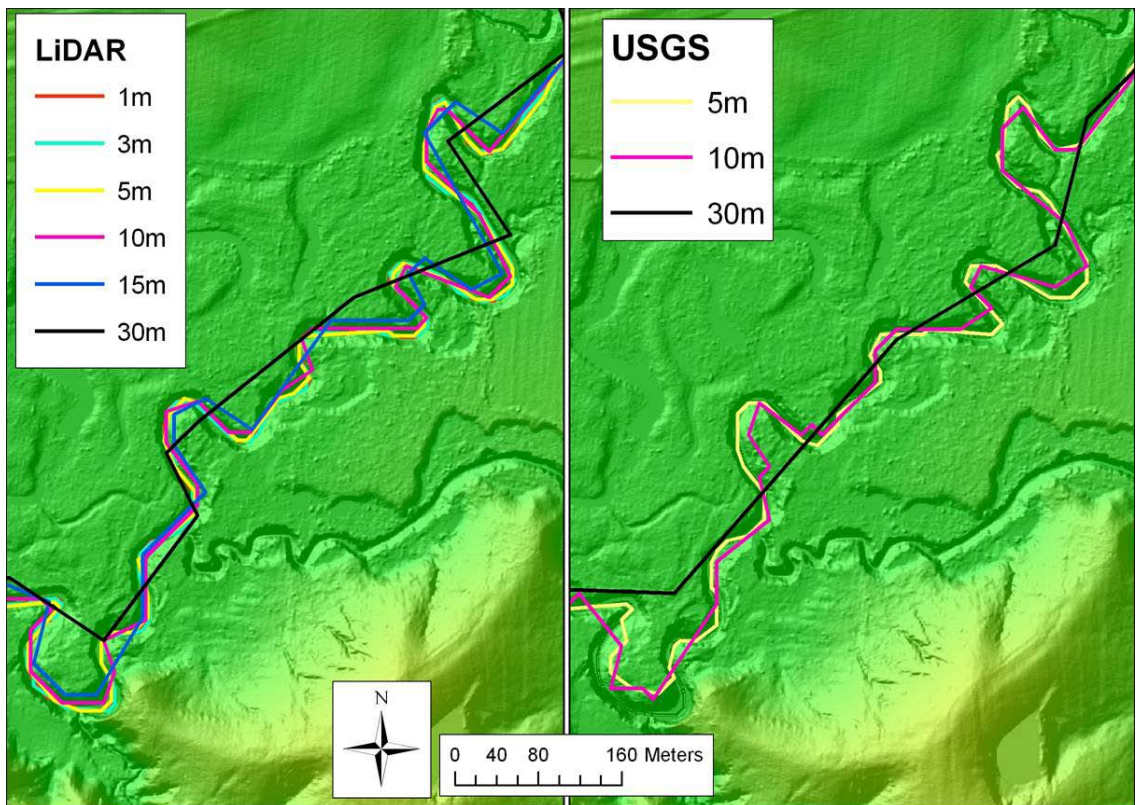

Figure 24: Stream flow paths

The APCA tool was used to quantify the offset of the flow paths using the

LiDAR one-meter derived flow path as the reference. Flow paths all meandered

across each other several times (Figure 24). The nature of this flow path

delineation means the minimum distance apart for the linear features is always

zero.

Table 6 shows the discrepancies of derived flow paths compared to the

one-meter LiDAR flow path. The goodness of fit for these flow paths is also

normalized by DEM resolution to understand how the DEM plays a role in the

disparity. The maximum variation represents the farthest point along the flow

path away from the reference one-meter flow path. This point varied in location

for each resolution but showed a steady increase as the resolution decreased

(Figure 25).

Table 6: Stream spatial discrepancy (APCA - meters)
Proportion = variance / raster resolution

DEM	Min.	Max. / Prop.	Mean / Prop.	Std. Dev. / Prop.
LiDAR 1m	0	0 / 0	0 / 0	0 / 0
LiDAR 3m	0	3.61 / 1.20	0.74 / 0.25	0.69 / 0.23
LiDAR 5m	0	8.06 / 1.61	1.20 / 0.24	1.01 / 0.20
LiDAR 10m	0	35.85 / 3.59	3.39 / 0.34	3.28 / 0.33
LiDAR 15m	0	43.14 / 2.88	6.20 / 0.41	5.59 / 0.37
LiDAR 30m	0	76.01 / 2.53	13.36 / 0.46	11.92 / 0.40
USGS 5m	0	39.00 / 7.80	1.66 / 0.33	2.93 / 0.59
USGS 10m	0	90.35 / 9.04	5.11 / 0.51	8.69 / 0.87
USGS 30m	0	236.00 / 7.87	18.76 / 0.63	26.47 / 0.88
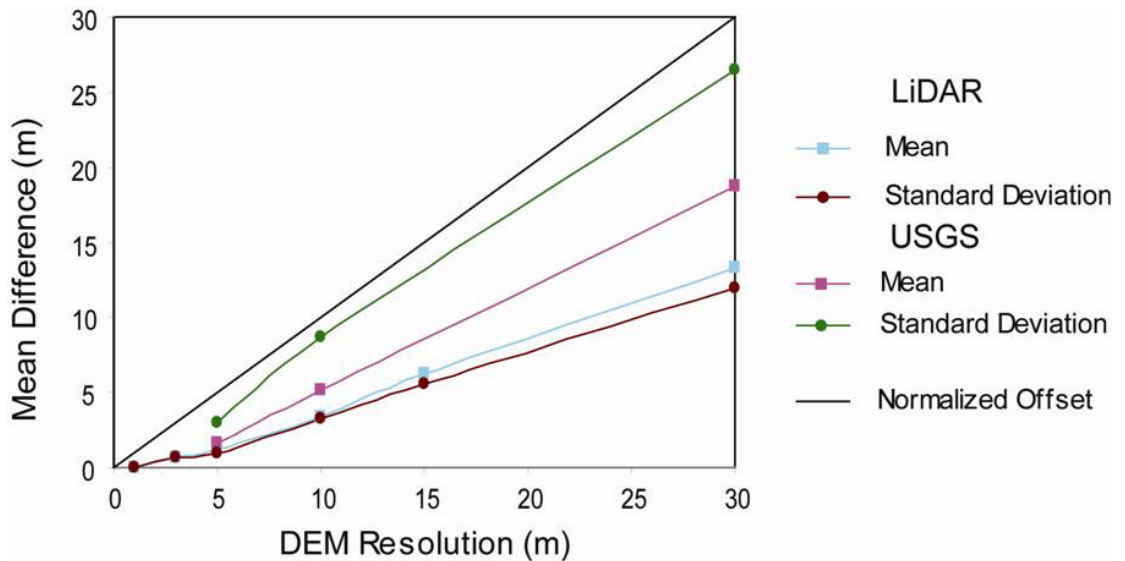
Figure 25: Mean stream offset
Normalized offset (Cell resolution = meters)

Normalizing these lines with regards to their cell size yielded surprising results. The ten-meter datasets (LiDAR and USGS) were the farthest normalized distance away from the reference with the averages tapering lower in both directions (Figure 26). When analyzing maximum offset, all results from LiDAR calculations resulted in less than one normalized cell offset. The USGS five-meter lines represented the only instances where the offset distance was greater than the normalized offset (Figure 26).

This was not the case when analyzing the mean difference. In both the mean distance and the normalized mean distance, as the resolution decreased, the offset of the lines increased. However, all mean differences of the LiDAR flow paths offset were less than one half cell resolution. The mean USGS flow paths followed the same patterns but the average offset was higher, reaching 0.63 percent cell size at thirty-meter DEM resolution (Figure 26).
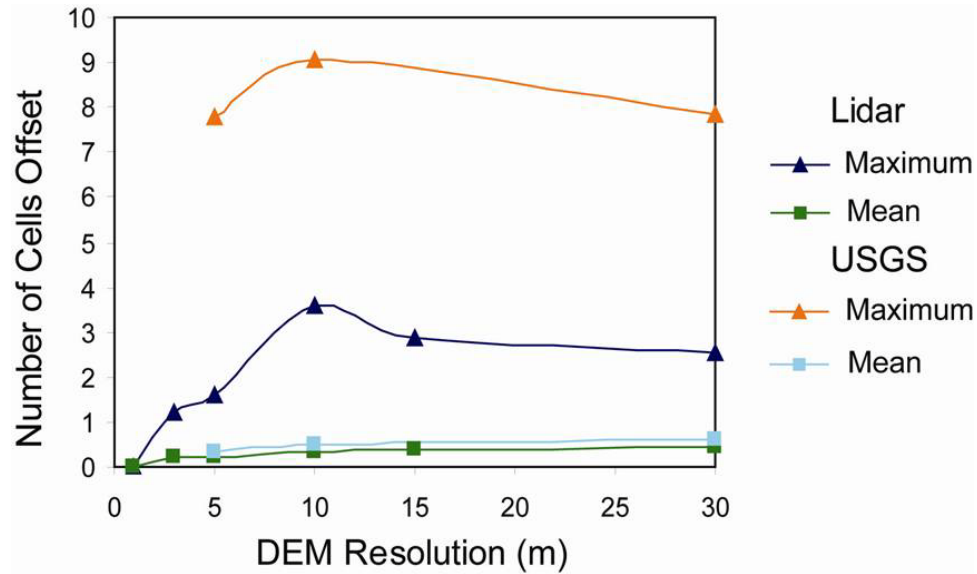
Figure 26: Offset proportions (Maximum and mean)
Normalized Offset (Cell resolution = meters)

Another variation between two data source flow paths involves the offset in standard deviation. In all cases the standard deviation increases as the resolution increases. This suggests that as the resolution increases, the range of variation found using the APCA tool includes a wider range of values. The variation is much greater when analyzing the USGS flow paths against the reference stream as the standard deviations for the 5-, 10-, and 30-meter flow paths had a range more than twice the comparable LiDAR flow path. The five-meter USGS flow path had a standard deviation 2.9 times as large as the five-meter LiDAR flow path.

5.2 Watershed Boundaries

The hydrologic model also delineates a drainage basin for all cross section locations. This basin represents all water that could contribute to flooding

at that location. The watersheds for this study ranged in size from just under 5.2

to just under 46.5 km^2 of drainage (Table 7). The data suggests that a

relationship exists between DEM resolution and the size of the watersheds they

produce. As the resolution of the DEM decreased, the calculations showed a

slight decrease to the area of the delineated watersheds. This trend was

consistent for both the LiDAR and the USGS DEM watershed boundaries (Figure

27).

Table 7: Watershed drainage area (km^2)

DEM	Maximum	Minimum	Mean
LiDAR 1m	46.41	5.18	25.23
LiDAR 3m	46.38	5.18	25.20
LiDAR 5m	46.33	5.13	25.17
LiDAR 10m	46.33	5.10	25.20
LiDAR 15m	46.10	5.13	25.15
LiDAR 30m	46.15	5.13	25.12
USGS 5m	46.26	5.07	25.15
USGS 10m	46.10	5.07	25.07
USGS 30m	46.15	5.07	25.04
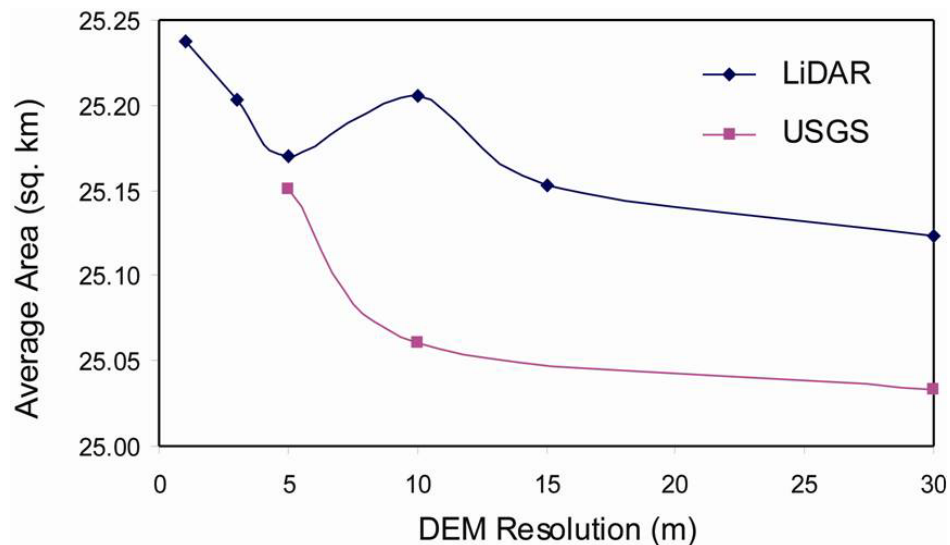
Figure 27: Average drainage basin area (km^2)

Higher resolution watersheds are going to be comprised of much longer boundary lines due to the variations within the cell sizes. While the total area difference between a 1- and a 30-meter watershed might be small, the location and total length of the boundary potentially varies greatly. To analyze these linear discrepancies, the APCA analysis was used (Figure 28).
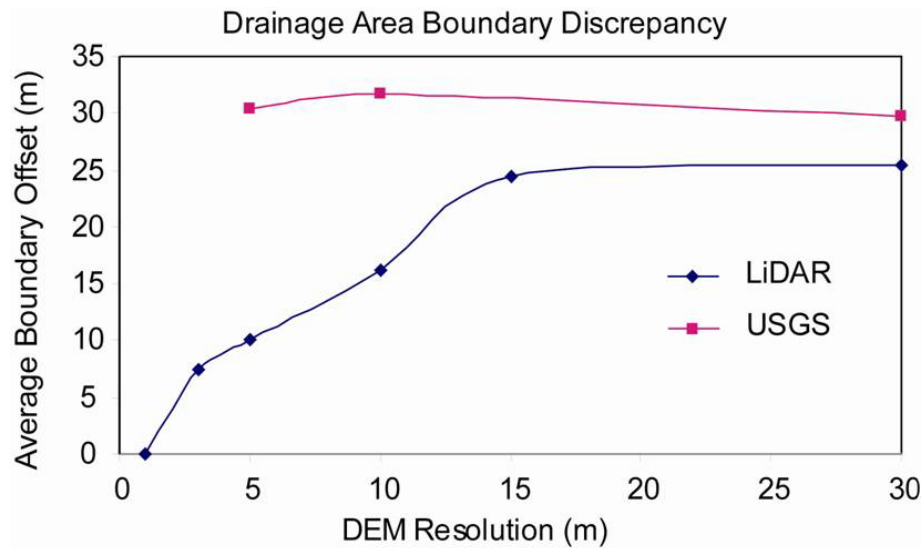
Figure 28: Watershed boundary discrepancy (Average)

One clear trend was observed that the lower resolution LiDAR DEMs produced boundaries that were on average farther away from the reference boundary. The LiDAR boundaries ranged from an offset distance of 7.4 meters to 25.4 meters on average. The USGS DEMs were more consistent in the amount of total offset distance with a range from 29.6 to 31.7 meters regardless of resolution, though this total distance was consistently higher than the LiDAR boundary.

The study area includes contributing areas that range from 5.2 to over 44 square kilometers of drainage. When these drainage basins were grouped by

basin area it was determined that the size of the boundary seemed to play no role in the total offset distance as determined using APCA (Figure 29). The basins area were grouped into four size ranges (<15, 15-26, 26-36, and 36+ km^2). The overall grouping results trended consistently, but there was no distinct correlation between drainage basin area size and amount of delineation discrepancy. This was consistent for both LiDAR and USGS DEM derived data sets.
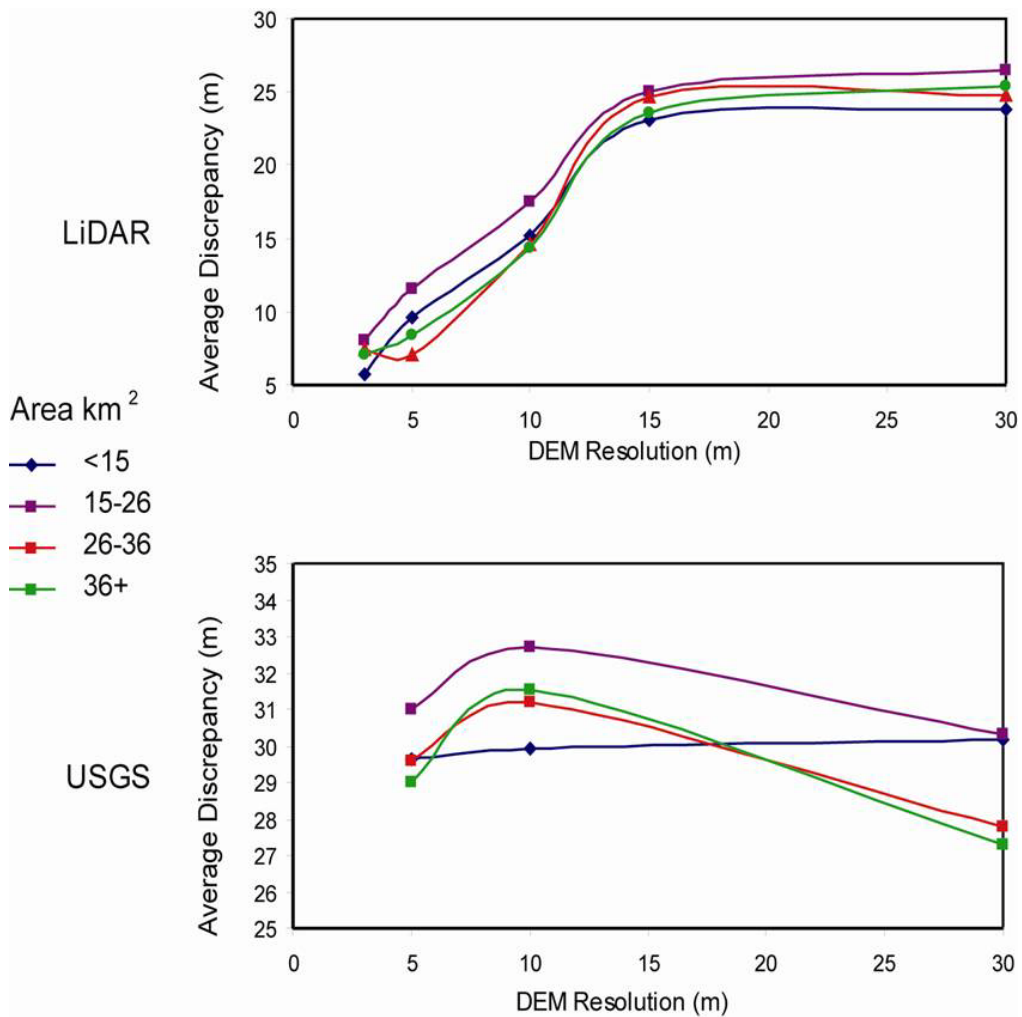
Figure 29: Watershed boundary offset discrepancy (Average by area)
APCA mean boundary offset distances from reference boundary

The boundary comparison allows for analysis to find outlier portions of a watershed that have been delineated. The linear fluctuations along the watershed boundaries tended to be similar with each DEM. The LiDAR watersheds also trended towards a better goodness of fit on average than the USGS watersheds, but this too might be a variation introduced due to different source data.

This analysis method was used to find possible errors within the models themselves. Figure 30 (A) shows a dramatic spike in discrepancy around the 18 square kilometer drainage area. Upon closer inspection the one-meter dataset was actually responsible for all other boundaries being offset by such a large distance. As the one-meter boundary was used as the reference dataset a common spike at all resolutions indicates a problem with the reference watershed.

A more disjointed spike appears at approximately 41 square kilometers of drainage area (B). This fluctuation shows that some type of variation occurred at that location resulting in a lower than average goodness of fit at five of the eight watersheds. This type of spike indicates a likely problem and further inspection of the cross section placement or DEM is needed.
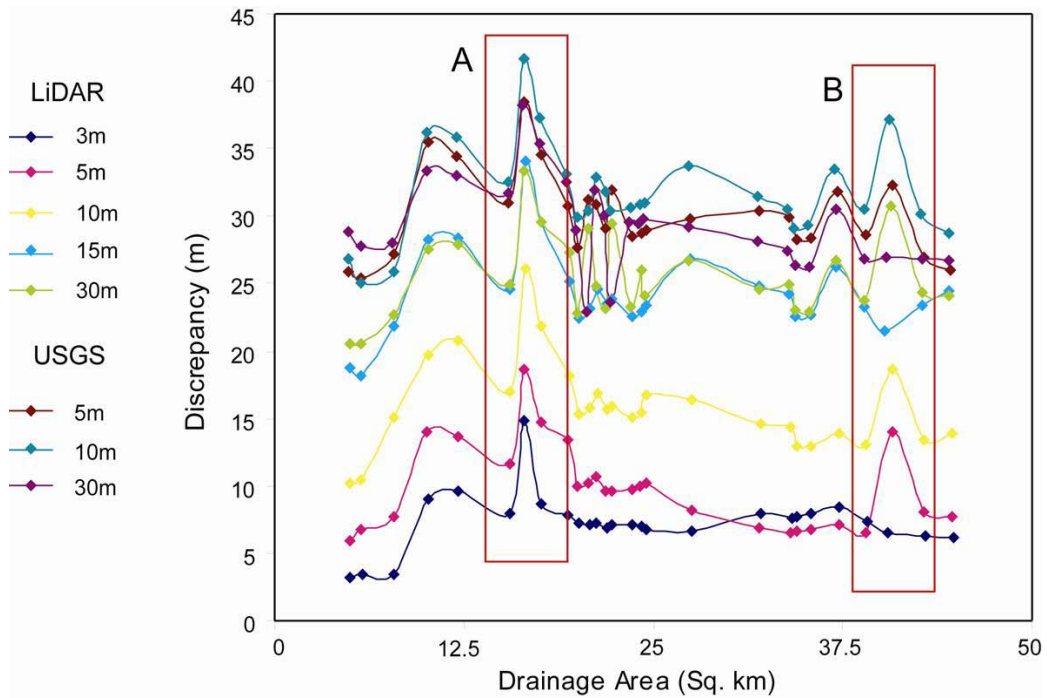
Figure 30: Boundary offset discrepancies by watershed (APCA mean)
Two types of discrepancy spikes in data observed (A / B)

Using the one-meter LiDAR watershed boundary as the reference introduced some discrepancy to the overall watershed analysis but these variations were found throughout the various resolutions. Most discrepancies were contained to individual cross section drainage basins, though some occurred along the outer ridge of the study area (Figure 31). These fluctuations introduced a measure of variation to all cross sections downstream from that location.
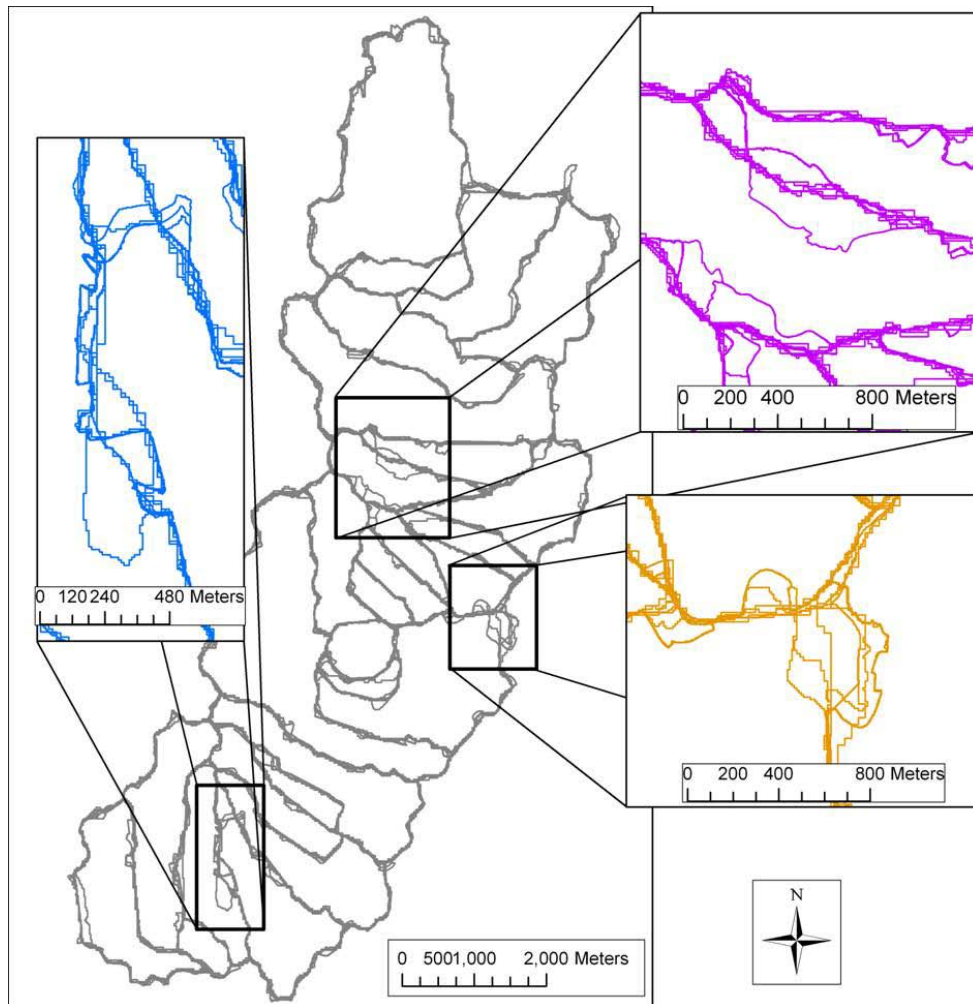
Figure 31: Boundary discrepancies
Varying DEM resolutions cause variety of drainage basin boundary errors
Single basin errors / multiple basin errors / errors on all basin boundaries

The major reason for these discrepancies is that watershed boundaries delineated by a one-meter LiDAR DEM are susceptible to minor terrain variations. Figure 32 takes a closer examination of the large spike in variation that occurs at approximately 18 kilometers of drainage. The red line representing the reference boundary (1-meter LiDAR) clearly moves south of the ridgeline when it encounters a terrace. This error is further complicated when the boundary encounters the road running north and south: it follows the road across the ridgeline before it begins to correct itself well north of the ridgeline. At the

75

same location the three- and five-meter LiDAR boundaries leave the ridgeline, veering north along a separate terrace. This area compounds the errors that produce a lower goodness of fit for those two boundaries at that location.
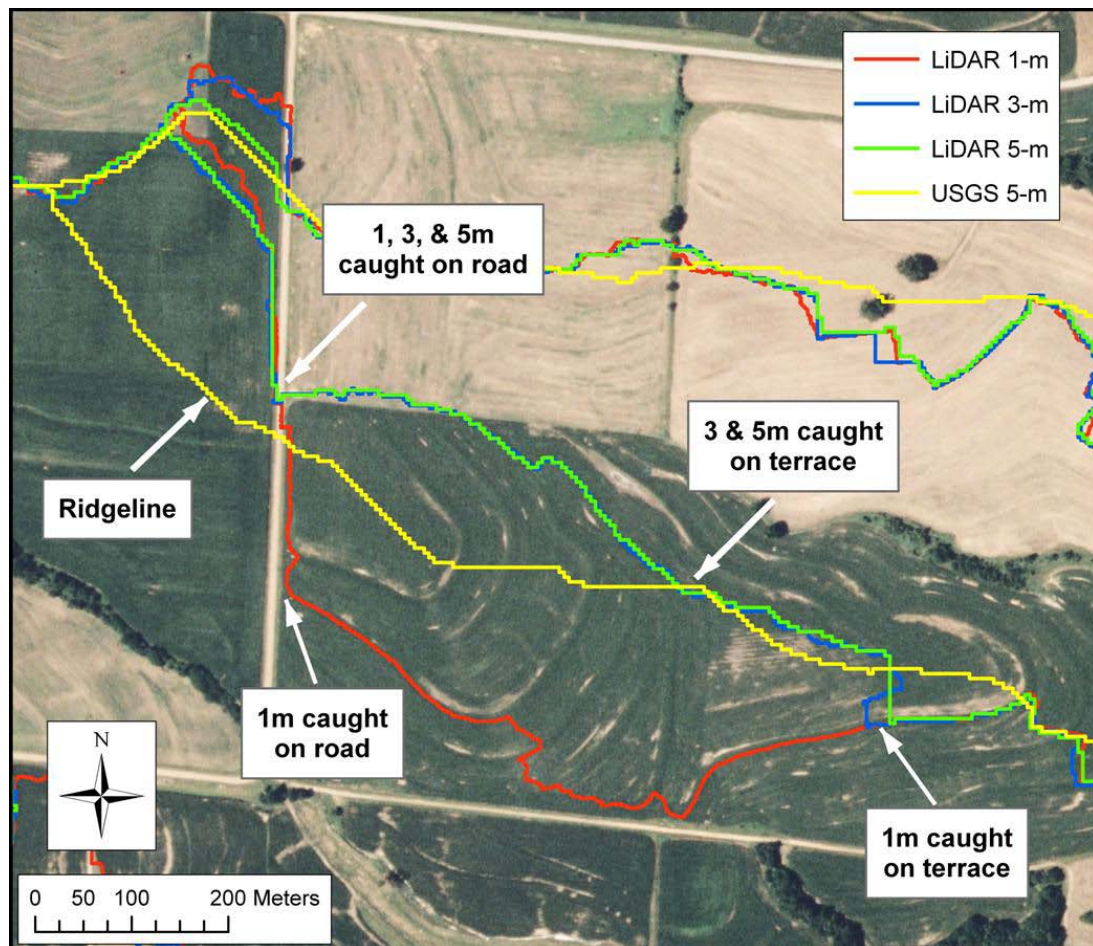
Figure 32: Boundary discrepancy example
Variations in drainage basin delineation using higher resolution LiDAR

The drainage basin boundary discrepancies found in Figure 32 can be attributed to how the one-meter watersheds were derived. The one-meter watershed boundary showed a tendency to follow features other than the ridge tops. While the one-meter boundary was used as the reference for all calculations, it clearly introduced some disparity to the results.

5.3 Hydrologic Parameters

The cumulative effect of the DEM resolution changes begins to become apparent as the hydraulic model input variables vary. The input data for HEC-RAS consists of the stream flow path, cross section locations, discharge, energy grade slope, Manning's "n" value, and bank positions. Stream flow paths created from higher resolution DEMs having more vertexes increases the amount of data input into HEC-RAS and affects the resultant computations. Additionally, the length has a small effect on the slope calculations imported into HEC-RAS.

The CARES analysis model calculates all its equations in standard units. These calculations are converted to metric equivalents upon input to the HEC-RAS hydraulic model therefore slope results discussed here are reported in standard units. Table 7 shows the average computed slope change at all cross section locations in both 'feet per feet' (local slope) and 'feet per mile' (watershed slope). Variations in flow path length introduced variations to the points where elevations were extracted from the DEM thus varying the elevations used to calculate the slope. The slopes were also normalized using each cross section's drainage area to see if the size of the watersheds played a role in any calculated slope variations. Each cross sectional slope was normalized by its corresponding drainage basin area prior to being averaged allows for slopes from varying watershed sizes to be compared directly.

There were 27 cross section locations used for modeling the study area across a wide range of drainage areas. To understand how the slope calculations might be affected, each slope was normalized by the area of the

watershed they represented. Table 8 indicates that the normalized slope

calculations show the overall trend of higher calculated slopes as the DEM

resolution gets smaller. The results also indicate a possible data error with the

LiDAR 10-meter calculations as the normalized results for that DEM defy the

trend of the other LiDAR calculations but the difference is not significant to the

average 'feet-per-mile' calculation that was to be input into HEC-RAS.

Table 8: Slope calculations

DEM	Watershed Slope (feet/mile)	Local Slope (feet/feet)	Normalized Watershed Slope (Area)
LiDAR 1m	16.08	.0082	2.60
LiDAR 3m	16.10	.0081	2.59
LiDAR 5m	16.74	.0087	2.71
LiDAR 10m	16.72	.0102	2.62
LiDAR 15m	17.69	.0090	2.79
LiDAR 30m	18.43	.0085	2.87
USGS 5m	15.18	.0103	2.36
USGS 10m	16.16	.0079	2.57
USGS 30m	18.61	.0118	2.86

Another input value needed for HEC-RAS to operate is the discharge.

The resultant portion of the regression equation is symbolized by the letter Q

while the number represents a specific return interval (100 = 1% annual chance

flood hazard). By calculating a Q100 discharge at every cross section location,

some model interpolation is removed during the HEC-RAS modeling process.

The Q100 representing the discharge for each cross section location shows

some variation when average discharge is calculated. When the Q100 is

normalized by drainage area, a very similar pattern is derived (Figure 33).

Both methods of analyzing the calculations show that there is a distinct increase in calculated Q100 as the DEM resolution decreases. As Q100 is calculated using both watershed area and watershed slope, both are susceptible to variations introduced during the calculations.
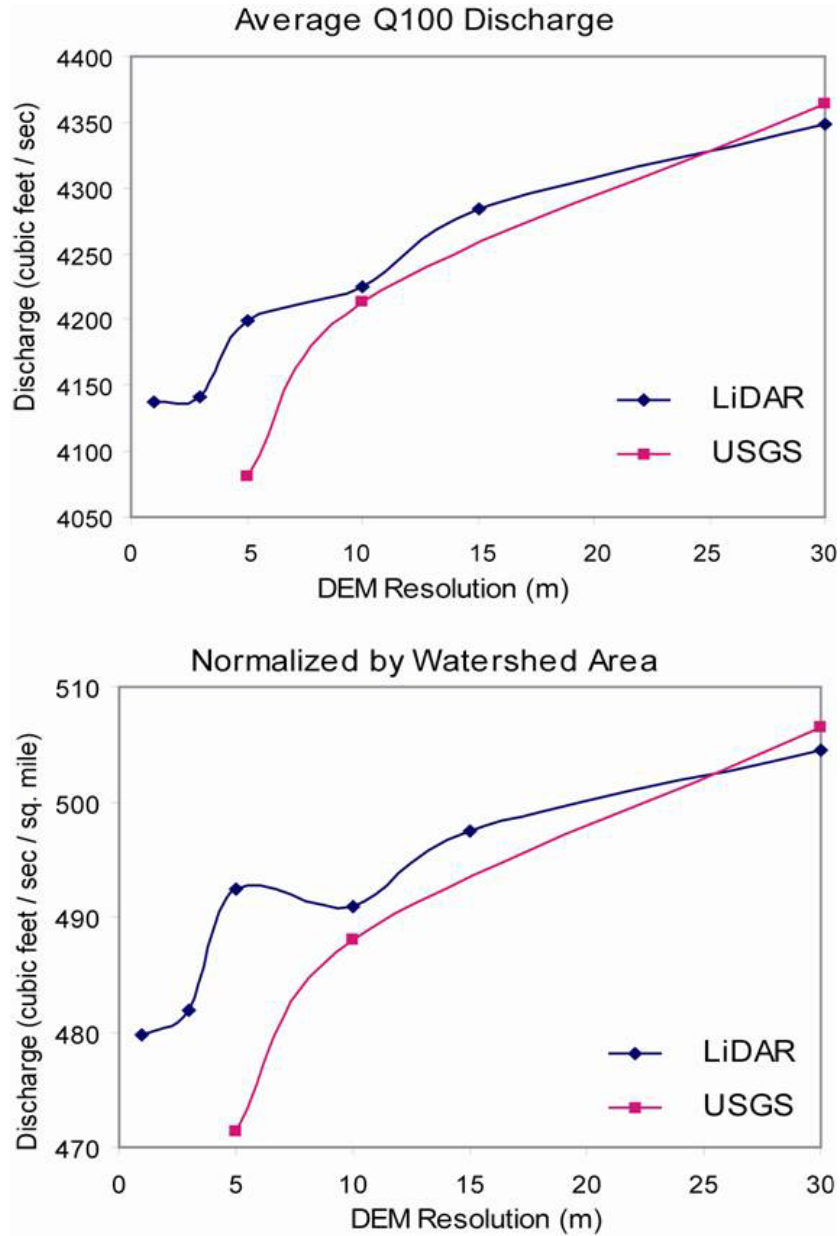
Figure 33: Q100 Values
Calculated Q values for 1% annual chance event by DEM resolution
Calculated Q values for 1% annual chance event by DEM resolution / Area

5.4 Flood Water Heights (Water Surface Elevations)

All calculations from HEC-RAS were exported back to their original GIS cross section layer. The CARES model performs both hydrologic and hydraulic calculations so a comparison of values between two types of models is possible. The two models used the same values to calculate their flood elevation levels through the two separate methods. Some variation is expected when comparing HEC-RAS calculations with the CARES model as HEC-RAS adjusts the water surface based upon the interactions with all cross sections. When all corresponding cross sections and resolutions results were compared, the CARES model calculated water surfaces are lower than the HEC-RAS results by a net mean difference of -0.1 meter (Figure 34). The overall trend of the difference results in a statistical bell curve pattern, suggesting a normal distribution of elevation difference.
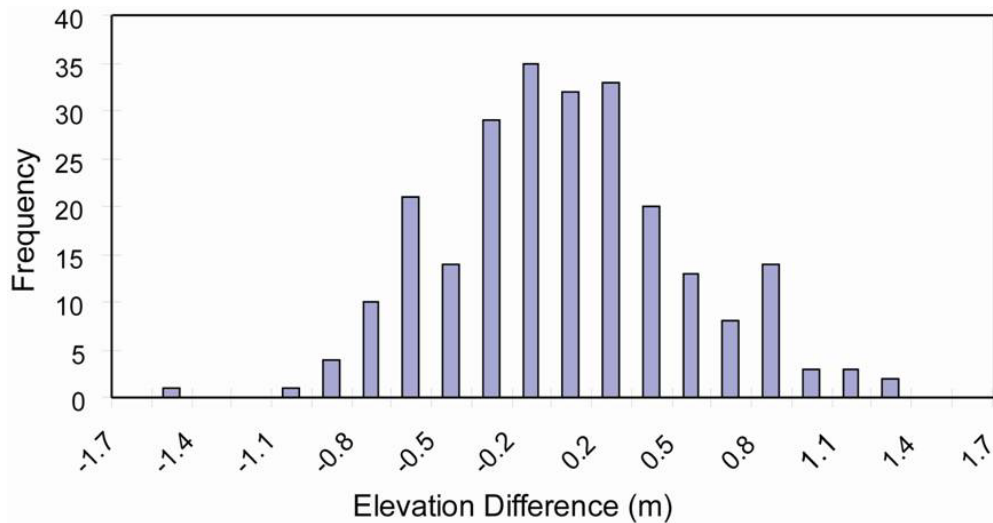
Figure 34: Distribution of water surface calculation discrepancies
CARES model – HEC-RAS model water surface elevations
Total average offset CARES model water surface -0.1m on average

When these calculations were broken down by DEM resolution no patterns were apparent (Table 9). Each cross section location and each resolution (243 water surface calculations per model) were compared by subtracting the HEC-RAS calculation from the CARES water surface. The USGS five-meter resolution shows the largest range of variation at 2.67 meters, but that result is skewed by a single cross section that calculates at 1.55 meter difference between the two models.

DEM	Mean	Low	High	Range
LiDAR 1m	-0.17	-0.82	0.69	1.51
LiDAR 3m	-0.05	-1.03	0.78	1.81
LiDAR 5m	-0.00	-0.77	1.03	1.80
LiDAR 10m	-0.14	-0.74	0.65	1.39
LiDAR 15m	-0.08	-0.82	0.81	1.63
LiDAR 30m	0.06	-0.50	0.73	1.23
USGS 5m	-0.37	-1.55	1.12	2.67
USGS 10m	-0.24	-1.00	0.29	1.29
USGS 30m	0.10	-0.63	1.20	1.83
All	-0.10	-1.55	1.20	2.75

Table 9: Comparison of modeled elevations (m)

When comparing these results by DEM resolution the 5- and 10-meter USGS DEMs introduced the largest variation between the model calculations on average. Conversely, on average the LiDAR 5-meter DEM produced an average offset of zero. As Figure 34 shows, the majority of the cross sections displayed an average discrepancy of less than 0.6-meter difference in elevation (higher or lower) regardless of the DEM used.

5.5 Floodplain Delineation

The floodplains delineated were varied both in location and formation. The total area covered by delineated floodplains shows some surprising results. When floodplains were modeled, the cumulative effect of resolution changes tended to increase the overall area. The floodplains generated using LiDAR data showed a small increase in overall size for the study area of .15 square kilometers. When the secondary floodplains were generated using the common cross section elevations, the lower resolution actually decreased the overall size of the floodplains by .18 square kilometers.

The USGS DEM analysis shows some of the same pattern as the original floodplains increased .28 square kilometers between the five-meter and thirty-meter floodplains. While all three USGS floodplains decreased in size, the thirty-meter DEM delineated floodplain did not show the same proportion of decrease in size as the LiDAR floodplain (Table 10).

Table 10: Modeled floodplain area (km^2)

DEM	Original Analysis	Secondary Analysis	Change
LiDAR 1m	2.49	2.49	0
LiDAR 3m	2.54	2.41	-0.13
LiDAR 5m	2.56	2.41	-0.15
LiDAR 10m	2.54	2.38	-0.16
LiDAR 15m	2.64	2.35	-0.29
LiDAR 30m	2.64	2.31	-0.33
USGS 5m	2.15	2.07	-0.08
USGS 10m	2.25	2.02	-0.23
USGS 30m	2.43	2.35	-0.08

In order to examine the boundary offset, floodplain polygons were converted into line features and analyzed using the APCA tool. Each bank was analyzed individually using the LiDAR one meter floodplain as the reference dataset (Figure 35). This same linear feature was used as the reference during the secondary analysis using consistent elevations as the one-meter floodplain was identical for both analyses.
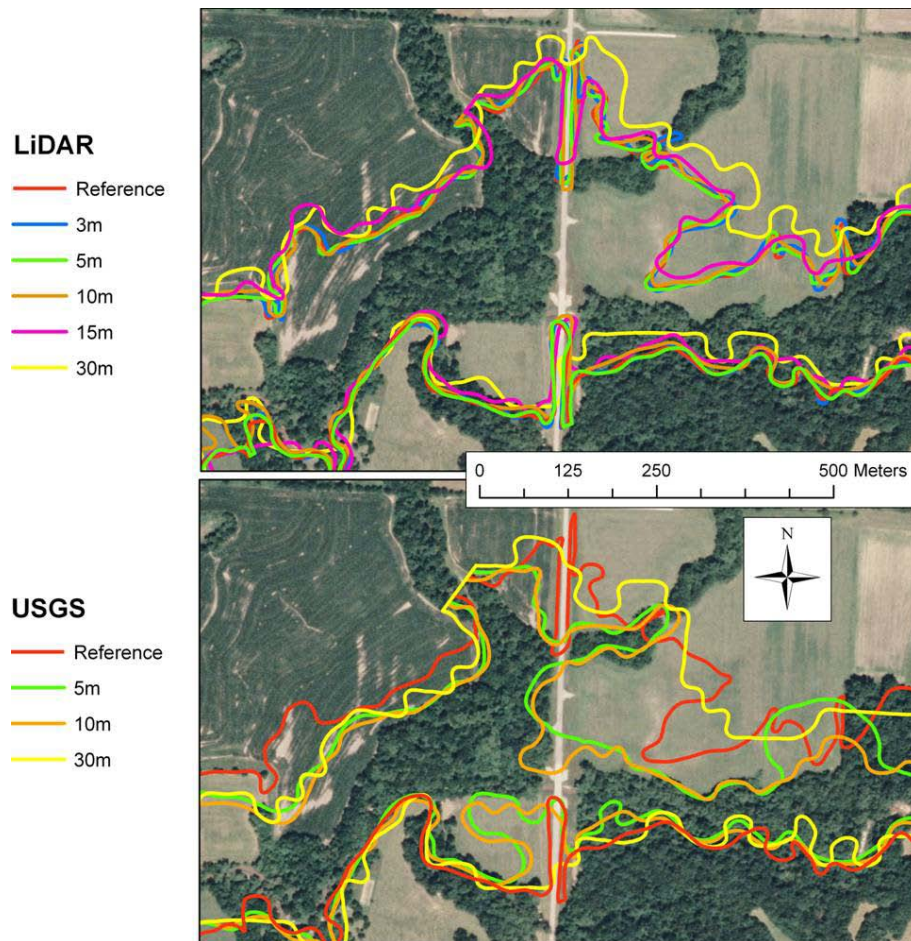
Figure 35: Floodplain boundaries at highway E
Edges of floodplains showed variation due to resolution of DEM

The results from the APCA analysis presented a mean offset between the modeled floodplain bank and the one-meter LiDAR floodplain bank. This offset

represented a measurable distance that showed consistent increase as the

resolution decreased. This was consistent with both the original LiDAR

floodplains as well as the secondary set of floodplains using the common

flooding elevations. The USGS DEMs tended to have approximately the same

average distance discrepancy regardless of which resolution or set of floodplain

elevations was used. All the USGS DEMs also had much higher levels of

absolute discrepancy than the LiDAR floodplains (Table 11).

Table 11: Spatial discrepancy (meters)
Proportion = variance / raster resolution

DEM	Right Max / Proportion	Right Mean / Proportion	Left Max / Proportion	Left Mean / Proportion
Primary Analysis				
LiDAR 1m	0 / 0	0 / 0	0 / 0	0 / 0
LiDAR 3m	103.0 / 34.3	4.8 / 1.59	56.8 / 18.9	3.6 / 1.18
LiDAR 5m	121.6 / 24.3	7.2 / 1.43	72.1 / 14.4	5.7 / 1.15
LiDAR 10m	109.6 / 11.0	9.9 / 0.99	119.4 / 11.9	8.2 / 0.82
LiDAR 15m	126.6 / 8.4	15.1 / 1.01	104.7 / 7.0	11.3 / 0.75
LiDAR 30m	138.8 / 4.6	18.7 / 0.62	108.0 / 3.6	15.7 / 0.52
USGS 5m	124.0 / 24.8	25.9 / 5.17	112.7 / 22.5	17.4 / 3.47
USGS 10m	123.4 / 12.3	27.1 / 2.71	100.2 / 10.0	17.3 / 1.73
USGS 30m	145.3 / 4.8	26.1 / 0.87	99.0 / 3.3	18.1 / 0.60
Secondary Analysis				
LiDAR 1m	0 / 0	0 / 0	0 / 0	0 / 0
LiDAR 3m	78.3 / 26.1	3.8 / 1.28	58.7 / 19.6	3.1 / 1.02
LiDAR 5m	83.2 / 16.6	4.8 / 0.95	55.1 / 11.0	3.7 / 0.73
LiDAR 10m	76.1 / 7.6	7.1 / 0.71	72.0 / 7.2	5.9 / 0.59
LiDAR 15m	86.4 / 5.8	10.6 / 0.71	69.4 / 4.6	9.2 / 0.61
LiDAR 30m	112.4 / 3.8	16.9 / 0.56	92.3 / 3.1	14.3 / 0.48
USGS 5m	157.6 / 31.5	27.2 / 5.43	100.2 / 20.0	16.1 / 3.22
USGS 10m	144.0 / 14.4	28.1 / 2.81	117.0 / 11.7	17.6 / 1.76
USGS 30m	145.0 / 4.8	24.5 / 0.82	98.2 / 3.3	17.3 / 0.58

While the secondary round of floodplains produced using common elevations did not seem to affect the USGS discrepancies, the removal of cumulative modeling changes created a noticeable difference within the all LiDAR models. The differences shown in Figure 36 between the primary and secondary delineations can be attributed solely to changing the resolution of the DEM. Standardizing the floodplain elevations prior to delineation resulted in up to a 36 percent increase in goodness of fit.
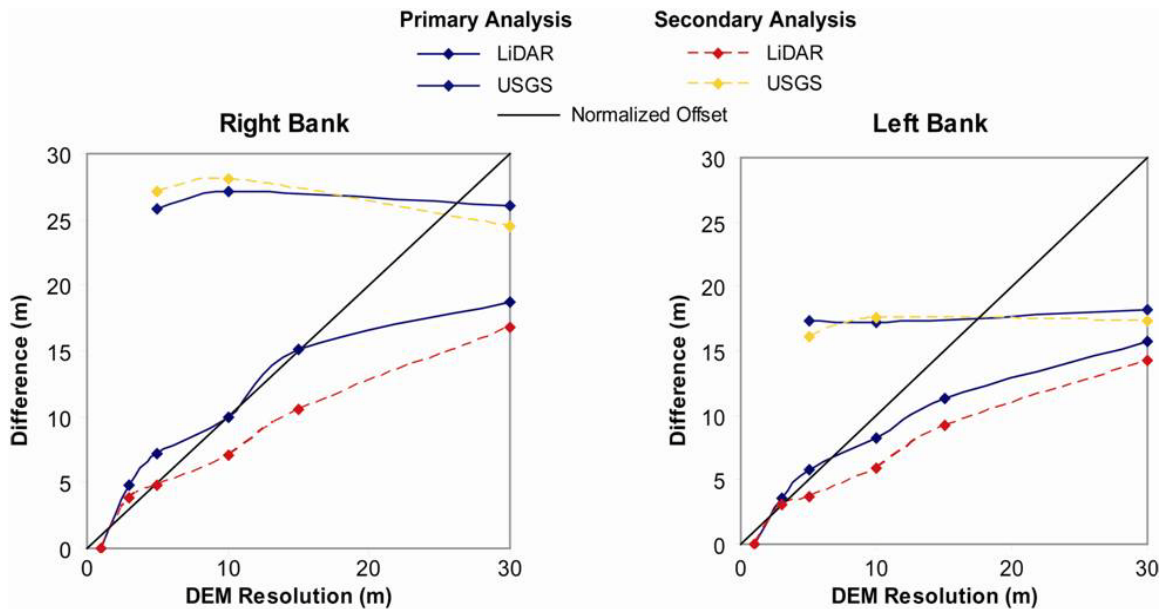
Figure 36: Absolute goodness of fit (Average distance for banks)
Comparison of initial calculations and secondary removal of cumulative variation

The amount of discrepancy found for each bank was normalized against its DEM resolution to understand how the cell size affected the results. The normalized analysis presents a similar trend where the USGS floodplains have a much lower goodness of fit (Figure 37). The five-meter USGS DEM delineated floodplain was at best a minimum of three cells away from the master floodplain

on average. Using the same analysis on only the LiDAR DEM floodplains

produced much closer matches: no floodplain averaged more than two cells

away from the master.
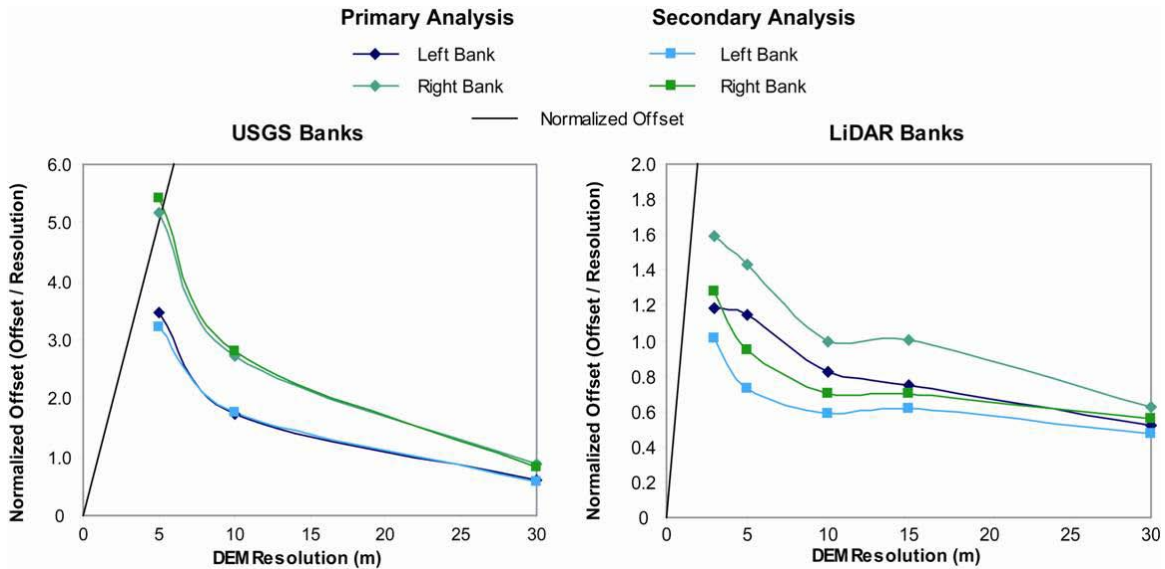
Figure 37: Relative goodness of fit for banks (Normalized by resolution)
Comparison of initial calculations and secondary removal of cumulative variation

Additionally the secondary analysis of floodplains produced dramatically

better offset results because it was designed to eliminate much of the cumulative

effect of the modeling process. This secondary analysis used the elevations that

produced the reference floodplain (LiDAR one-meter data) and by removing the

discrepancies introduced during the hydrologic and hydraulic processes

floodplains were produced that were quantitatively closer to the reference

floodplain.

The analysis of the offset distance was consistent in that as resolution

increased, the offset distance ratio to the cell size decreased. At five meters, the

left bank of the ratio produced by USGS DEMs was just over one normalized

offset (greater than one cell size in distance). Under all other conditions, the

offset distance stayed within one resolution. In all cases, once the cell size

reached thirty-meter resolution the boundaries fell within a one-cell discrepancy

(Figure 37).

The overlap fit analysis was then used to determine the goodness of fit for

the entire flooded areas (Table 12). The flooded polygons were converted to

raster datasets with one-meter resolution for common analysis. The use of the

overlap fit calculation resulted in a distinct linear relationship between the

resolution and the fit percentage for the LiDAR floodplains. Each floodplain was

compared to the LiDAR one-meter floodplain and as the resolutions got smaller,

the fit percentage decreased (Figure 38).

Table 12: Overlap fit percentage calculations

DEM	Primary Analysis		Secondary Analysis	
	Overlap Area (km^2)	Fit %	Overlap Area (km^2)	Fit %
LiDAR 1m	2473	100	2473	100
LiDAR 3m	2419	93.9	2408	94.7
LiDAR 5m	2405	91.6	2413	94.5
LiDAR 10m	2355	88.7	2391	91.1
LiDAR 15m	2330	83.7	2356	86.6
LiDAR 30m	2236	77.9	2380	79.5
USGS 5m	1889	68.9	2060	72.1
USGS 10m	1945	70.3	2033	71.6
USGS 30m	2059	72.3	2351	70.0

The USGS floodplains actually showed contradictory results. During the

initial analysis the fit percentage increased as the resolution decreased, though

the entire range showed a change of less than five percent. During the

secondary analysis using the non-cumulative elevations to create floodplains, the

USGS floodplains reversed to agree with the LiDAR results where decreased

resolution meant lower fit percentage. The range on this analysis was a

difference of just over two percent, suggesting that resolution had little effect on

the USGS fit percent calculations (Figure 38).
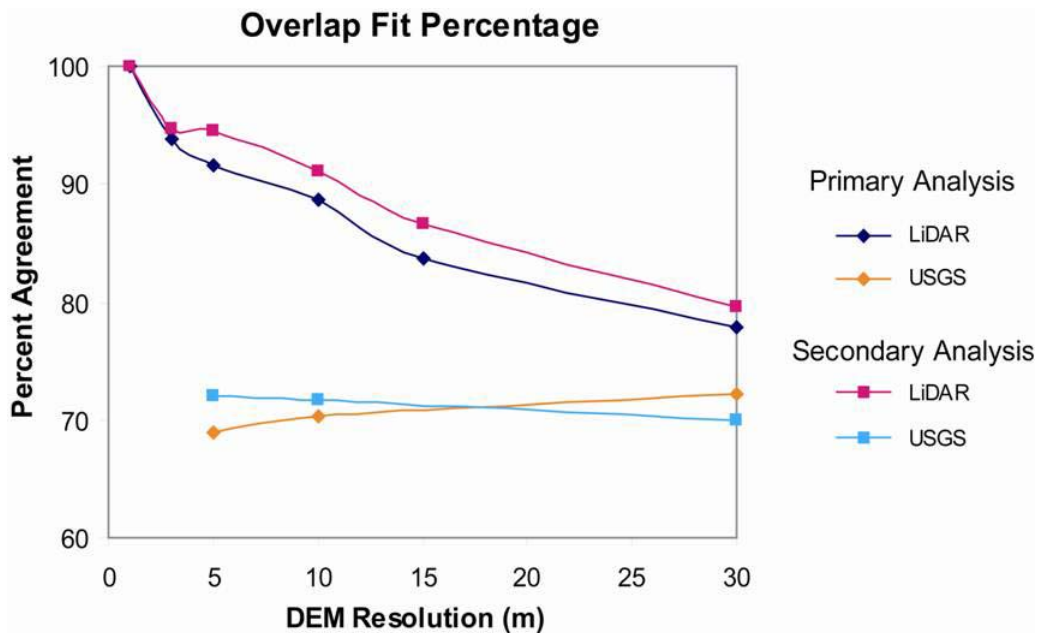
Figure 38: Overlap fit percentages
Comparison of initial calculations and secondary removal of cumulative variation

The final method of analyzing the floodplains used a floodplain probability

index. This index ranking of one-meter squares identifies the locations that were

attributed as being located within a delineated floodplain. Cells that were not

delineated within any floodplain were not attributed, therefore a 100 percent

agreement would indicate all flooded cells matched within all modeled

floodplains. The primary calculations using all DEM resolutions results in a 48

percent agreement between all floodplains. Substituting the secondary elevation

modeled floodplains does increase the agreement, but only to 52 percent

agreement. There are quite large percentages that fall under eight (89 percent),

seven (78 percent), or one (11 percent) floodplain under these scenarios. An

index of one indicates outlying fingers of flooded areas that were accounted for

by one variation of the DEM, whereas the seven and eight represent areas that

were derived most commonly as flooded but were missed by only one or two

DEMs (Figure 39).
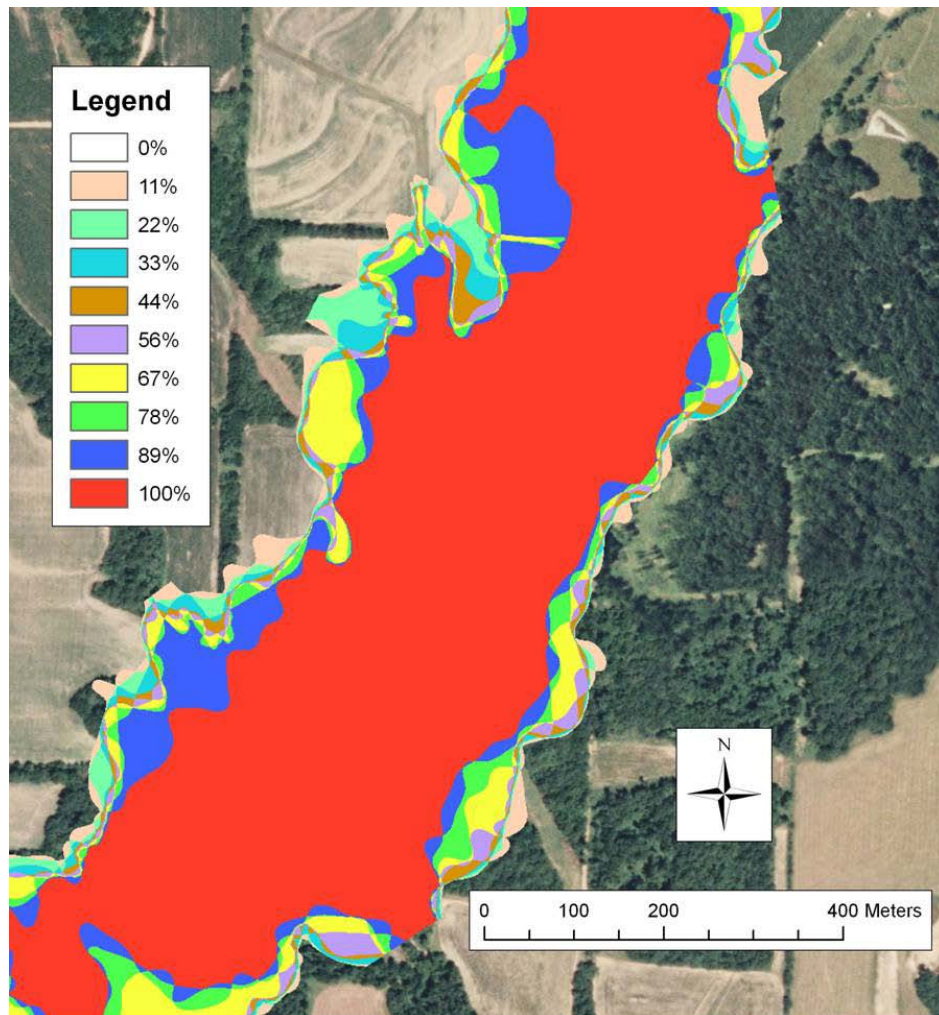
Figure 39: Comparative flooding index
Percentage of DEM resolutions indicated flooding in a cell location

The floodplains produced seemed to be more affected by the different source data between the LiDAR and USGS DEMs. When the same probability index was applied to only the LiDAR DEM floodplains a much higher overlapping resulted. A rating of six represents a perfect overlap of all LiDAR floodplains and the initial model runs resulted in 68 percent agreement. When adjusting for cumulative model variations during the secondary calculation the agreement percentage increased to 73 percent (Figure 40).

The resulting index figures suggest that the cumulative variations introduced through the modeling procedure account for approximately five percent variation between floodplains. The original nine floodplains increased four percentage points while the LiDAR-only runs increased five percent when using the same elevations for modeling. The standardization of elevations also tended to reduce the number of outlying cells that were only modeled by one floodplain.

Overall the results indicate that using variable resolutions of DEMs produces floodplains of multiple size and position. For purposes of delineating floodplains the highest resolution tended to create the most meandering floodplains. As shown using the overlap fit percentage analysis, it was found that using a re-sampled DEM to create floodplains uses less resources and time, and can produce similar results (Figure 41). As these results are based on using a modeled reference floodplain, this should be tested against known flooding events to determine if there is a similar result pattern.
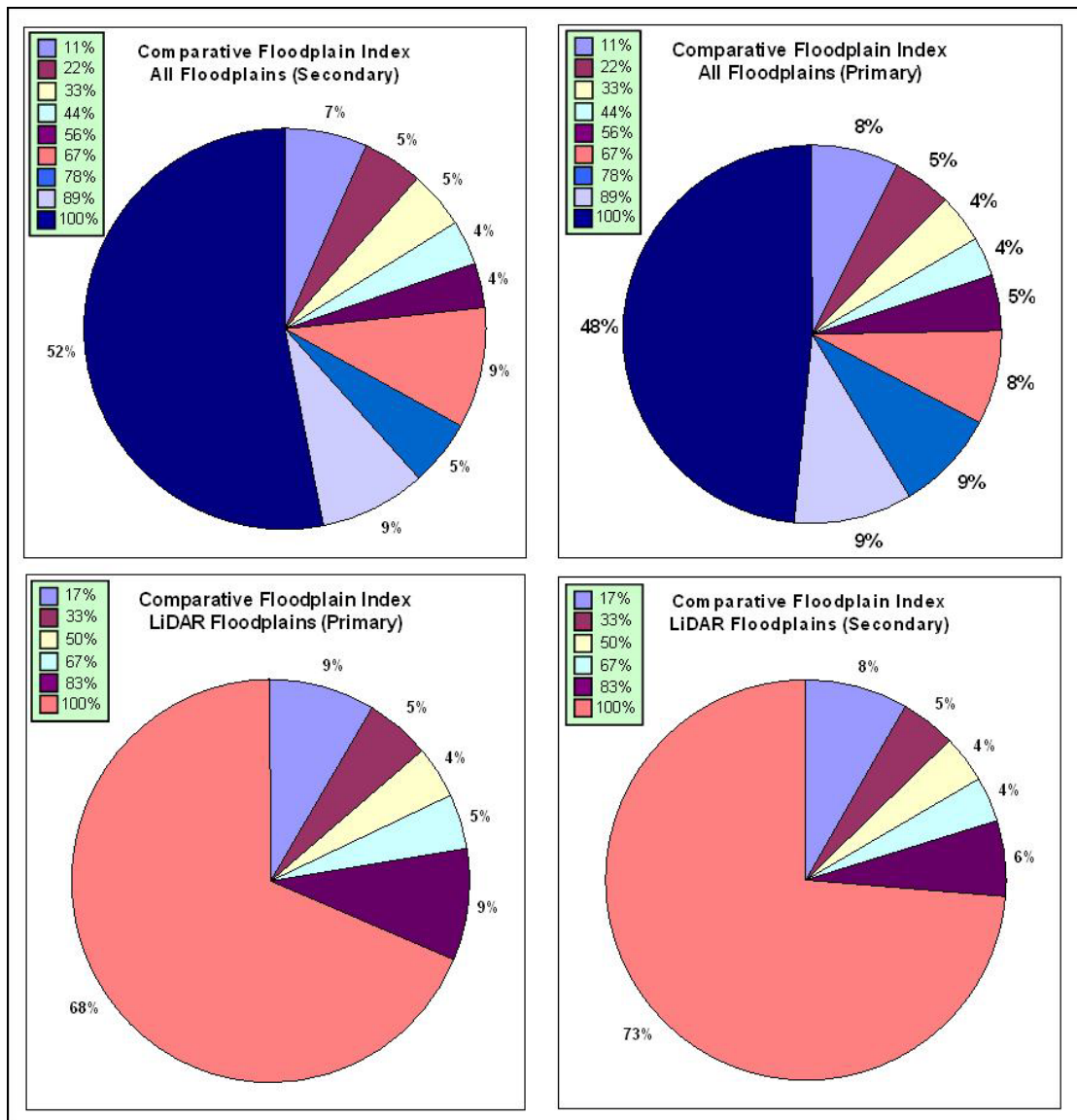
Figure 40: Comparative flooding index results
Comparison of all resolution calculations and only LiDAR DEM calculations
**Comparison of initial calculations and secondary removal of cumulative variation**

Original modeling using a three-meter DEM produced results within 5 percent of the one-meter data when combining boundary and area analysis. If the cumulative effect of the model resolution variation is eliminated, both the three- and five-meter DEM floodplains varied by less than 5 percent. Overall the discrepancy between base data was much larger as the USGS floodplains

91

tended to be approximately 25 percent different regardless of resolution, although using this method it is impossible to say which set was most accurate when compared to actual flooding events.
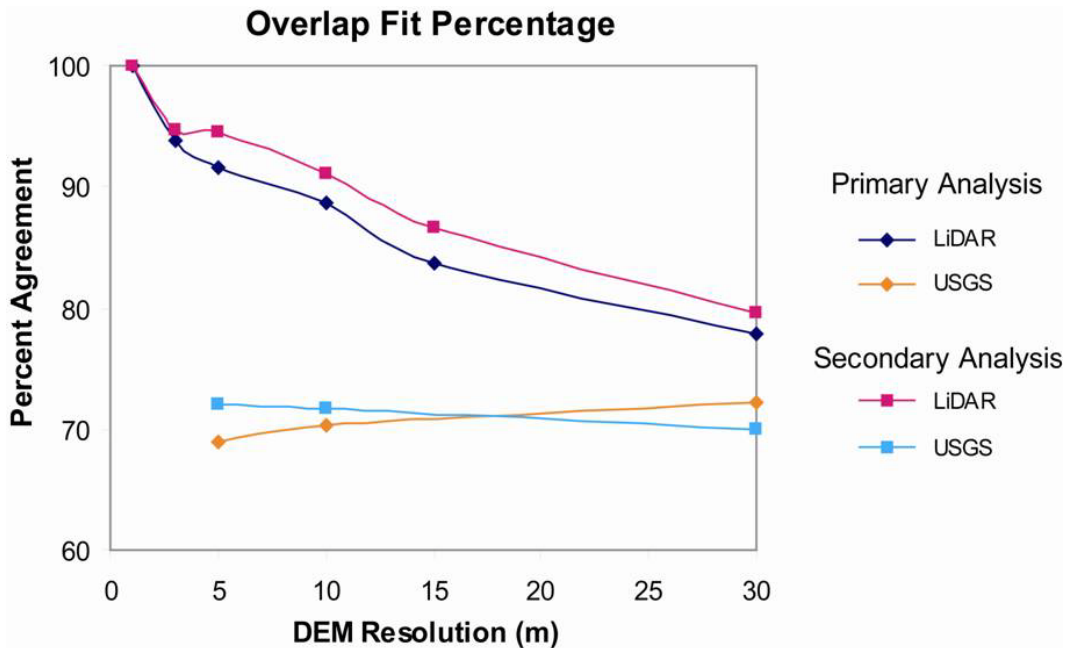
Figure 41: Overall variation percentage from reference floodplain
Comparison of initial calculations and secondary removal of cumulative variation

The usage of re-sampled DEMs for production of floodplains can serve as a substitute when dealing with restrictions on drive space, computing power, or time (Table 12). Using the models described above, the entire modeling process using a one-meter DEM for all steps created a floodplain in just under 24 hours while the same programs took a little over 2.5 hours to calculate data at thirty-meter resolution. Substantial savings of calculation time using these models was seen during the flow path delineation process where multiple grids were created and the hydrologic process where cross section drainage basins were

delineated. These modeling processes relied heavily on interpreting or manipulating the grid data while other portions of the process did show the same trending of file sizes and time savings, but to a lesser degree.

This study ran several processes using AML programs that factored into the time line. The hydraulic modeling portion of the process using HEC-RAS did not produce any extra data. Results were imported back and forth from the same cross section coverage while the HEC-RAS project files were all approximately the same size and so were not included in Table 13. Additionally, because all analysis grids were created at one-meter resolution, the file size and time savings of the analysis portion of this study were all similar.

Table 13: Model Time and File Size (LiDAR data)

DEM	Flow Path Delineation[1]			Hydrologic Modeling[2]		Hydraulic Modeling[3]	Delineation Modeling[4]	
	Size (kb)	Time (min)	Size (mb)	Time (min)	Size (mb)	Time (min)	Time (min)	Size (mb)
1m	435000	420	835.0	870	52.8	45	100	1.88
3m	44400	115	93.7	410	32.4	42	75	1.83
5m	16000	60	33.7	200	9.9	38	60	1.76
10m	4020	35	8.5	170	5.0	35	50	1.58
15m	1790	20	3.8	145	3.9	32	35	1.51
30m	475	10	1.0	90	2.3	30	25	1.46

1 Flow Direction Grid, Flow Accumulation Grid, Flow Path Cover
2 Cross Section Cover, Watershed Covers, Profile Covers
3 Time needed to generate Input/Output files
4 Dams Cover, Floodplain Cover

The study area was approximately 46.6 square kilometers covering a single stream. Flooding generally occurs over much larger areas, requiring a much larger dataset. Riverine flooding generally develops over longer periods of

time than flash flooding or levee breaches, and thus can give a modeler lead time

to analyze flood potential provided updated elevation models exist.

Chapter 6

Discussion

This study investigates just one portion of the floodplain modeling process. The process itself has many variables and this study's focus is singularly on the effect of DEM resolution on floodplain delineation. Analysis of modeled floodplains is important for disaster preparation, insurance regulations, and protection of life. Automated floodplain generation is a complex process that requires multiple models and techniques, along with more than a little human interaction. Even though it is impossible to completely mimic or predict flooding events, determining the best source data and methodology is imperative to produce the best possible floodplains. An accurate floodplain prediction could possibly mean the difference between losing everything owned to a flood, or even life and death.

The policies of the professionals who produce floodplain boundaries tend to determine what the most appropriate resolution would be for a study. Floodplain modeling for FEMA requires making a larger, more generalized floodplain to identify all properties that might need to purchase floodplain insurance, while individual municipalities need highly accurate floodplains for specific events. While in all cases more detailed methods (including field work) will produce a more accurate floodplain, the expense increases dramatically when compared to automated modeling. Additionally, even detailed flood modeling methods represent an educated guess when determining flooding

95

events as no two flood events are identical. It is with this in mind that this study provides a baseline for future modeling analysis.

This study area is chosen for its lack of impervious area, providing a good baseline of results that can be compared with future study sites with more impervious areas. Flooding in rural areas can be devastating, but in reality flood events in urban areas are of most concern to flood management agencies. Further study will need to focus on areas where flooding has greater impact, preferably areas with known flooding events allowing for comparison between reality and modeled floodplains. This study provides three analysis methods for comparison of flooding events that would prove useful in future studies of other areas and types of flooding events.

The study also only implements one model choice for each portion of the delineation process. This choice of model inserts a bias into the study that should also be investigated. Using different modeling techniques will result in different results, but the extent of how these differences affect the resultant floodplains should be analyzed. The choice of models for this study represents just one methodology approved by FEMA for the State of Missouri to develop floodplains that adhere to FEMA guidelines. Additionally, flooding models used for alternate purposes, or using different requirements might produce different results.

6.1 Stream Network Flow Paths

The analysis suggests that resolution has not only a profound effect on individual portions of the floodplain delineation process, but also has a cumulative effect. Resolution affects the creation of stream flow paths for modeling. Hydrologic models rely on these flow paths to calculate flood variables such as discharge and slope, and the results from this analysis suggest that a higher resolution DEM produces a more realistic flow path. The flow path created by one-meter DEM is almost an identical length to the original break line stream network developed with the DEM and matched aerial photography. Using a LiDAR DEM with a resolution of three meters produces a match of 99 percent while a five-meter resolution produces a flow path whose length is within the 97th percentile (Figure 42). As stream network delineation is the most time intensive portion of the floodplain modeling process (Table 12), it might prove useful when dealing with time sensitive analyses to use lower resolution stream networks that are relatively similar.

The results suggest that for this study area, the USGS DEMs with resolution of up to ten meters also produces a matching stream length within the 97th percentile. Caution should be used, however, when comparing DEMs derived from different source data. The ten-meter USGS flow path length is only 2.6 percent off from the referenced one-meter LiDAR DEM, while the ten-meter LiDAR length is 5.6 percent shorter. Individual analysis would need to be conducted on larger study areas before any substantive correlation could be claimed between data from different sources. The length for flow paths between

the USGS and LiDAR analyses tend to be quite similar but this should be

analyzed further with regards to location.
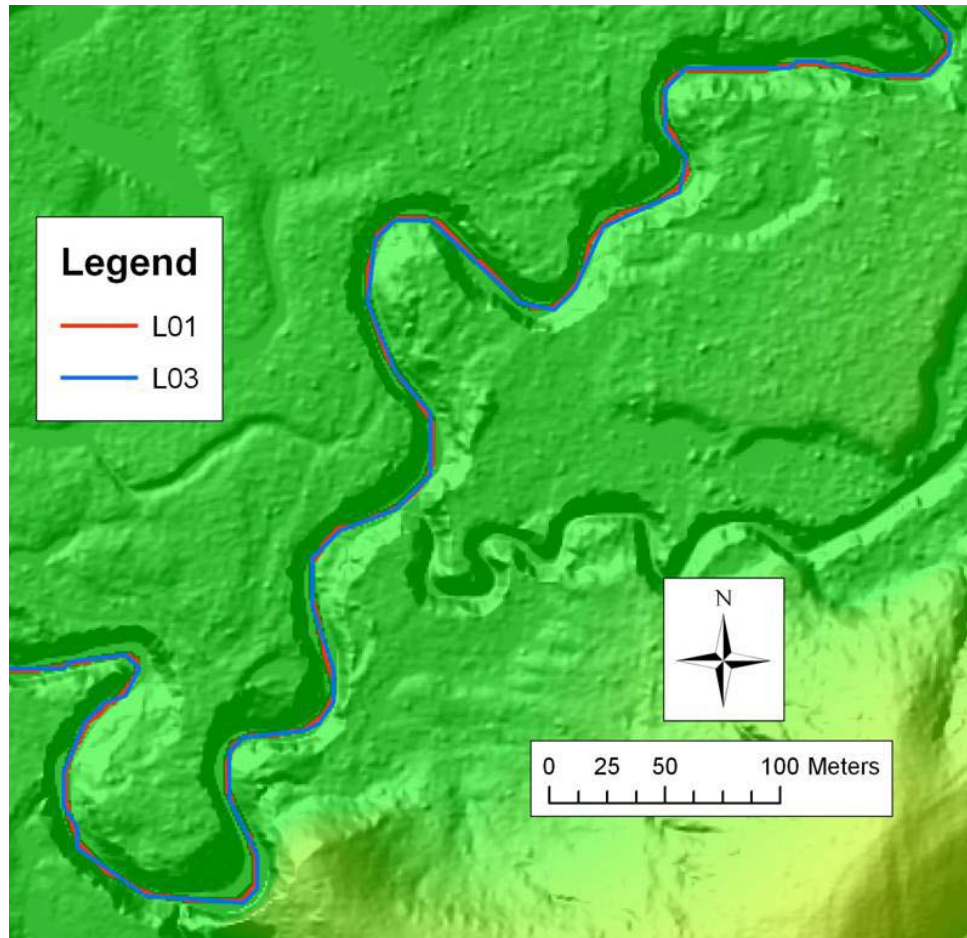
Figure 42: 1m and 3m LiDAR flow paths (97% agreement)

When looking at how these flow paths match the reference stream

network geographically, it appears that the resulting length calculations are

deceptive to some degree. The LiDAR flow paths at three and five meters both

produced flow paths that stayed within 1.2 meters on average of the master

stream network with the largest variation being just over eight meters. In both of

these cases the flow paths stayed within one quarter of a cell. When compared

to the reference, both the lengths and the locations of the three- and five-meter flow paths are very close, suggesting that not much accuracy is lost when re-sampling a one-meter DEM to five-meter resolution (Table 14).

Table 14: Re-sampled DEM flow path similarities

DEM	Length (km)	Length Difference (%)	Location Offset (m)	Location Offset (Cells)
LiDAR 1m	18.33	0.0	0 / 0	0
LiDAR 3m	18.12	-1.1	0.74	0.25
LiDAR 5m	17.83	-2.7	1.20	0.24

The USGS five-meter flow path is also close, with an average of 1.66-meter offset but a maximum offset of 39 meters suggests a large discrepancy in at least one location between the extracted and the reference flow paths. Overall the five-meter USGS flow path stays relatively close, within one-third of a cell, and when combined with the similarity in length this suggests that the USGS five-meter DEM in this location also serves as a suitable substitute. While the ten-meter USGS DEM had remarkably similar length to the master, the flow path location shows quite a bit of variation with an average of over five-meter offset and a maximum offset of over ninety meters.

The results indicate that the variation between a thirty- and one-meter DEM results in quite a large discrepancy between both length and location of flow paths. The 30-meter DEM available nationwide produces a flow path 20 percent shorter and has an average offset distance of 18.76 meters. The state of Missouri has a ten-meter DEM that produces a flow path with a mean offset of 5.1 meters. Some states such as North Carolina have a statewide LiDAR DEM

available. Even with LiDAR data though, there might be some types of models

that do not require perfect accuracy and can benefit from faster processing times

without losing much accuracy by resampling LiDAR to five-meter resolution.

6.2 Cross Section Drainage Basins

The thesis results suggest that delineated watershed boundaries get

slightly smaller when derived for floodplain calculations on average as DEM

resolution decreases. The area change from a one- to a thirty-meter drainage

basin decreases by .01 square kilometers on average. The variation in total area

is dependent on each particular cross section with some resulting in little-to-no

change while others vary more depending upon the surface ridges. Higher

resolution DEMs tend to produce less accurate boundaries under certain

circumstances. The discrepancies found in basins delineated from higher

resolution DEMs both added areas to basins that should not be there, and

subtracted some areas from basins (Figure 43).

In areas with terraces along hillsides, or areas around road surfaces,

these boundaries tend to move away from ridge tops based on these features.

This was more apparent when working with the LiDAR DEMs as these features

were described in the original one-meter data. In this thesis, the LiDAR results

show these variations in the one-, three-, and five-meter study. The data

sampling on the ten-meter DEM smoothed these areas enough to make the

features blend into the background and so the ten-meter LiDAR boundary follows

the ridge line more closely. The USGS DEMs do not have this problem as the

source data used to create them never had enough detail to show terraces, so even the five-meter USGS boundary follows the ridge lines accurately.
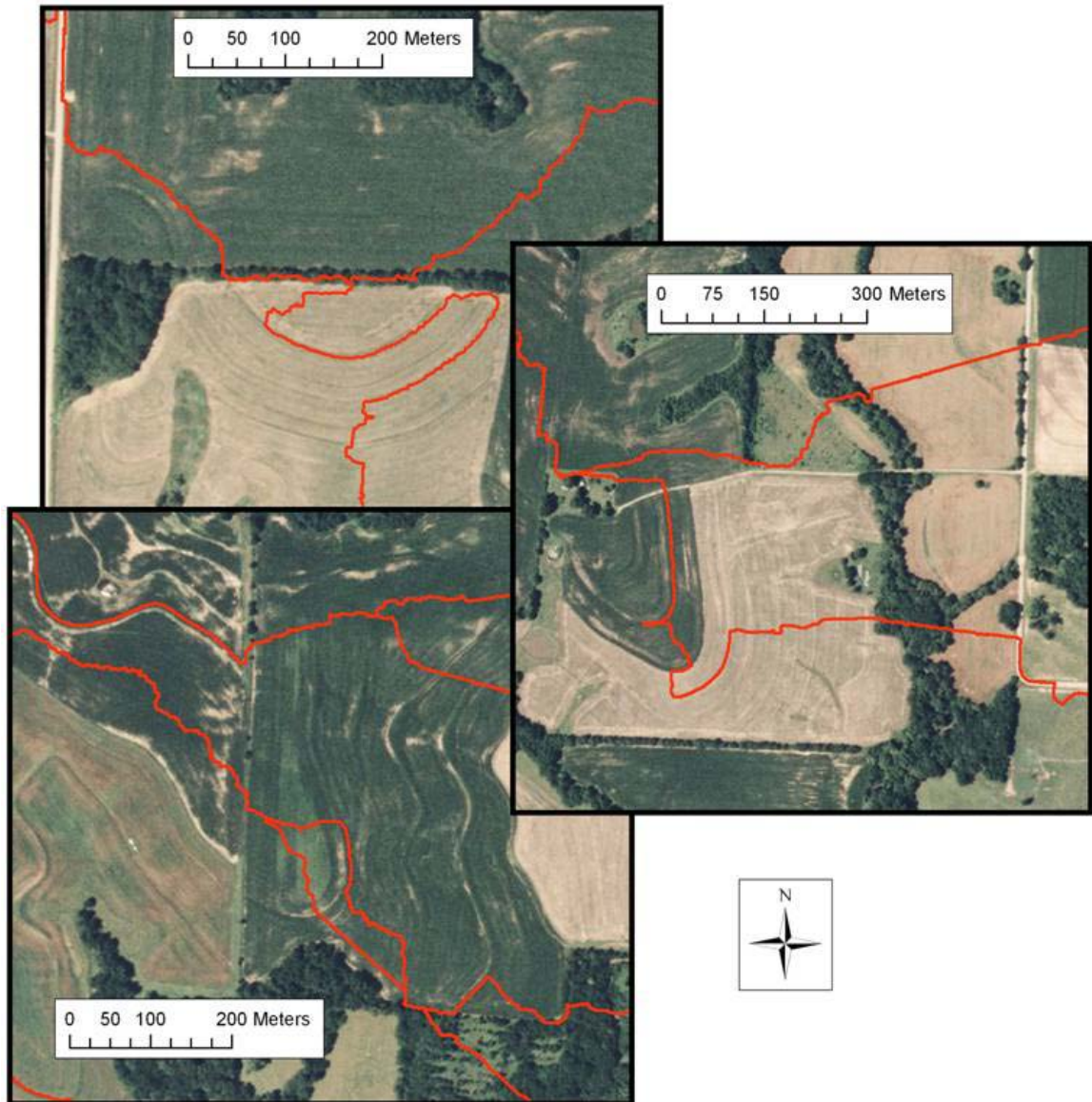
Figure 43: LiDAR watershed boundaries affected by terraces, roads, and other features Boundary lines should follow ridgelines but can be erroneous with LiDAR accuracy

When delineating watershed boundaries for area calculations, it appears that there is a trade-off between the different resolutions. The higher resolution DEMs tend to produce boundaries that follow all ridge meanders more closely,

101

creating a more accurate representation of the ridge line in most cases but occasionally losing the ridgeline entirely. The lower resolutions produce blocky watershed boundaries that are generalized in many places and that also lost the ridgeline in places.   This study found that either the ten-meter LiDAR or the five- and ten-meter USGS DEMs produce the most accurate watershed boundary. These were the highest resolutions that produced boundaries that follow ridge lines without being interrupted by minor elevation changes that introduced errors while minimizing the squaring of the boundary. The boundaries created by those resolutions tend to be marginally smaller in total area but the boundary is always centered on the ridge line.

One calculation for floodplain delineation is watershed slope. It is important that boundaries follow ridge lines to produce an accurate area for determining watershed slope. Lower resolution DEMs produce higher slope calculations over the same area suggesting that higher resolution DEMs produce the most accurate slope values. These higher resolution DEMs produce longer and more accurate stream networks; providing slope calculations with a more accurate distance between high and low elevations. This study's results suggest that one solution to this issue would be to use lower resolution DEM to generate cross sectional drainage area boundaries, and then apply those boundaries to the higher resolution DEM to extract elevations used for calculating slope and distance. Further study of this suggested outcome would be needed to determine the viability of such an interaction.

6.3 Floodplain Delineation

The cumulative effect of resolution changes on floodplain delineation proves integral to the floodplain production. Floodplains produced at lower resolutions tend to be larger in size regardless of source data. This is to be expected as lower resolution DEMs are more generalized due to cell size and provide less specific boundary blockages to contain flooding. The most surprising result from the study was that the USGS floodplains all tend to be smaller in size than the LiDAR floodplains. The largest floodplain modeled using the USGS thirty-meter DEM produced a floodplain .05 square kilometers smaller than the reference LiDAR floodplain. As the USGS DEMs are created using more generalized contour information it was originally hypothesized that they would produce larger flooding events.

Another surprising study result was that when the cumulative effect of resolution is removed from the calculated elevations during the secondary analysis, the floodplains produced trend smaller as the resolution decreased. There is no consistent change in elevations between the primary and secondary analysis to predict the reversal in trend direction that occurred with the LiDAR data. The USGS floodplains also decrease in size, though the decrease was not as consistent or linear and suggests that there is some calculation affecting the delineation process other than elevation.

The shape of the floodplains is studied by analyzing the boundaries, suggesting that lower resolution DEMs produce more generalized boundaries. Higher resolution boundaries tend to show more length due to increased number

of cell centers that provide vertexes but also due to the more detailed surface used for analysis. The results show a large discrepancy between the LiDAR and USGS produced boundaries suggesting large variations in the source data used to create these DEMs. Overall the changes within the LiDAR boundaries follows a consistent pattern, showing lower resolution DEMs create boundaries with more discrepancies while the USGS DEMs result in little to no change in overall discrepancy based upon resolution.

The general discrepancy between the LiDAR DEM and USGS DEM produced computations are probably due to the resolution of the original data used to create these two distinct DEM sets. The USGS DEMs are developed from USGS topographic 1:24,000 contour data. Similar original source data resolution produced USGS DEMs that result in similar model outputs regardless of DEM resolution because original data provides similar level of detail of the topography.

LiDAR DEMs are also derived from a single data source but with a much higher initial resolution. As the DEMs are re-sampled from higher to lower resolution cells begin to blend elevations, smoothing out more detailed information that is in the initial data. Therefore, model outputs are more resolution dependent. The best evidence presented is that watershed boundaries delineated using LiDAR DEMs follow some farm terraces and roads at certain scales while ignoring them at other resolutions.

Further analysis of the LiDAR boundaries shows that the secondary analysis tends to create much more similar floodplain boundaries. These

boundaries followed the same patterns of increasing with cell size, but the overall discrepancies drop by up to 34 percent (LiDAR five-meter right bank). These results allow for a measureable assigned to the amount of effect resolution changes using the DEM based CARES method for floodplain delineation. This measureable is not apparent with the USGS boundaries as the secondary analysis did not result in noticeable changes to the amount of discrepancy, possibly due to base data generalization during DEM creation.

The overlap fit analysis also provides a clear pattern when analyzing the LiDAR floodplains. These floodplains area follow a clear downward trend as the resolution is decreased. The 30-meter DEM produces a floodplain with a 78 percent overlapping fit percentage. The same trend is found when analyzing the floodplains produced during the secondary analysis even though all results are better overlap percentages than the primary analysis. This provides a methodology that compares calculated floodplains between resolutions, or could be adapted to compare modeled floodplains against known events. LiDAR provides the most accurate portrayal of the Earth's surface while a known flooding event could be derived from aerial photography. Based on the resolution of the aerial photography, the study provides a basic expected change that might be found. Floodplains produced using the USGS DEMs tend to be very similar in size and shape. Their best fit percentages are all very similar and unfortunately did not show a distinct pattern when compared to the LiDAR reference floodplain.

The floodplains that covered the same geographical location depend on the DEM source data to some degree; an area is modeled as flooded depending upon how the source data identified the elevations in an area. Variations between source data are important to the analysis results, controlling how well the floodplains match each other. Studying the entire area modeled "in the floodplain" by at least one model run provides a geographical analysis of how well these models agree, and conversely how much resolution changes these modeled floodplains. When analyzing the LiDAR floodplains, they tend to model the same locations within a floodplain 68 percent of the time, increasing to 73 percent during the secondary cumulative data analysis. Conversely there are 9 percent and 8 percent of these floodplains respectively that are outliers (Figure 39), only being modeled by one resolution.

When looking at all modeled runs, LiDAR and USGS combined, the percentages drop dramatically. Delineated floodplains agree in all instances only 48 percent of the time, increasing to 52 percent during the secondary analysis. This introduction of the USGS floodplains and their effect on the results suggests an underlying geographic difference between the source data that derived the two sets of DEMs. This is deceptive as the outliers actually accounted for less of the overall percentage at 8 percent and 7 percent respectively, but rather increases the number of cells that are 'almost' in agreement (Figure 39). The results suggest that the USGS floodplains are all similar with some variations that limit the number of included cells. This is probably due to the slightly smaller overall size of the USGS floodplain areas. The USGS delineated floodplains

tend to be smaller than the LiDAR floodplains so it follows that less area would be consistent between the flooded areas.

One implication of the study is that multiple resolutions should be used during automated floodplain modeling depending upon available source data. While using the highest available resolution always makes the most sense, should the chosen model call for calculating watershed boundaries, resampling the DEM cell size larger than 5 meters would seem appropriate. Using an original stream network that matches aerial photography would be the best course of action, though using a flow path generated from a DEM with 5 meter resolution loses very little actual length or horizontal accuracy. Floodplains themselves should be compared to known flooding events to discern the accuracy of the models where possible, although in general the LiDAR derived floodplains tended to model larger flooded areas.

Further study on the effect of resolution on floodplain delineation is needed. A multitude of variables affect the process and studying the process from a variety of different angles should provide a clearer analysis. The choice of models used in this study represents a basic limitation of studying floodplain modeling. The possible types of models for each portion of the process should be studied to determine its affect on model outputs. Only by cataloging various models, studying their inputs, outputs, and their overall impact on the floodplain modeling process can a determination be made as to the best available method. Future study of alternate models should also include a study of various DEM resolutions as inputs values.

Primarily this thesis determines that while LiDAR will produce the best possible representation of the Earth's surface, there are aspects of the floodplain modeling process that create more viable results using lower resolution elevation models. Additionally, it provides a measure of the expected difference for various study areas based upon the resolution of available DEMs. Using the highest resolution data available is always advised, this thesis provides a measure of just how much variation might be introduced into a delineated floodplain based upon its base resolution. The thesis recommends using LiDAR data at the natural one-meter resolution when time or data storage is not a concern, to model streams, calculate slope, and delineate floodplains, while resampling the LiDAR data above a five-meter resolution when calculating drainage basins is more appropriate.

6.4 Limitations and Further Study

This study is conclusive only for how resolution affects a basic agricultural watershed with limited impervious surface and slight slope characteristics. Additionally, the results are only indicative of specific models and their application to specific types of pre-existing DEMs. Conclusions drawn from this study should be carefully applied to other regions, data types, and models. Broad interpolation based upon the results from this study is not recommended until further study has been completed. This study area was chosen for its specific lack of defining characteristics so that it could serve as a control for future studies of flooding in other areas.

Future studies should be conducted to analyze these results in a multitude of ways. Studies should be conducted using various locations, land form and cover types. As flood modeling is done as a preventative measure against loss of property and life, primary focus should be on replicating the model in areas to study how impervious surface or human interruptions affect both flood modeling and underlying DEMs. Various locations would alter the study in many ways. Of primary concern for future studies would be analyzing flooding for more populated areas with more impervious surface, but would also look at variations such as elevation, slope, geologic formations, and various foliage effects, regression equations, among other factors.

While a very brief analysis of two types of hydraulic models' water surface calculations is a part of the study results, further study must be conducted to identify the most accurate modeling procedure. The results of this study suggested that both models that adapt water surfaces to include backwater and ones that strictly adhere to normal depth calculations produced very similar water surface elevations on average (within .15m). These were two hydraulic models, representing two of the hydraulic modeling families that could be chosen to delineate floodplains, and the type of terrain, land cover, and slope will alter those results. There are also many other models that could be chosen for hydrologic modeling or the delineation process. Additional study must be conducted to identify how well the results of other model combinations agree.

The primary focus of further study in any of these areas is to identify study areas that have known flooding events. Comparison of various land types,

model combinations, or any other floodplain modeling to known flooding events is the only method to ascertain the accuracy of the actual modeling process. Finding the model process that most closely represents flooding events or understanding that various models do better under certain land cover types or locations is important to all future study of automated floodplain modeling. It is also important to understand how variations such as land and plant cover, geologic formations, and slope are not only modeled but represented in the underlying DEMs.

LiDAR DEMs represent the best method to include all possible features using current technology so it becomes important to understand how these DEMs affect derived or modeled features such as floodplains. It is also important to determine whether or not representing all features is conducive to creating more or less accurate flooding events.   It is for all of these reasons that a plethora of future study areas are suggested to continue to analyze the variations of automated floodplain modeling begun by this thesis.

Chapter 7

Conclusion

The study looks at the effect of DEM resolution on the complex process of automated floodplain mapping. Using FEMA guidelines as the template, models are chosen to study how altering the resolution of the base DEM data underlying the calculations altered the outputs. The study area is the Camp Creek watershed in Saline County Missouri, a primarily agricultural watershed with slight slope and minimal impervious surface. The study area is analyzed using LiDAR DEMs re-sampled from one-meter data to 3-, 5-, 10-, 15-, and 30-meters. The same analysis is completed using existing USGS 5-, 10-, and 30-meter DEMs.

Hydrologic, hydraulic, and delineation modeling is completed for the study area nine times, alternating out the underlying DEM with each run. A series of nine stream flow paths, delineated cross section basins, and floodplains are analyzed. A secondary analysis using equalized water surface elevations as inputs for the delineation model create a total of 18 modeled floodplains to analyze within the study area. Analyzing the various stream flow paths, sets of cross sections, delineated cross section watersheds, floodplain bank locations, and floodplain polygons allows for conclusions about the similarities of each output. Linear features are analyzed using the APCA tool while areal features are analyzed using overlap fit percentage and a floodplain probability index.

The results of these analyses show that the LiDAR one-meter dataset produces data that is closest to reflecting the terrain in most cases. The resultant stream network and floodplains show meanders indicating a closer fit to the realities of the surface, while the meanders that are injected during the watershed delineation suggested a weakness of using the one-meter LiDAR DEM. It is determined that higher resolution DEMs tend to cause watershed boundaries to deviate from ridgelines.

Under the assumption that the LiDAR one-meter data would represent the most accurate flooding events, it is determined that re-sampling the LiDAR data to either three- or five-meter resolution for floodplain delineation results in large savings of computing time and data storage while producing results that match within five percent of the reference data. The thesis results suggest that it is more appropriate to use a re-sampled ten-meter LiDAR DEM to delineate watersheds as minor elevation changes in the highest resolution DEMs tend to cause errors when determining the ridgelines. This thesis suggests that a combination of DEM resolutions might prove most accurate when working in tandem during the modeling process.

The data suggest that the type of source data underlying a DEMs resolution also plays a key role in automated floodplain delineation. In this study, the USGS DEMs are all derivatives of the 1:24,000 contours and result in very similar calculations. The LiDAR DEMs all originated from a more accurate one-meter data set that is re-sampled to lower resolution, generally smoothing features away. Patterns are typically found in LiDAR data results that are not

112

present in the USGS data. This thesis attributes these effects to the LiDAR DEMs being derived from the higher resolution original data.

The results also suggest a large amount of cumulative error is introduced during the process depending upon the resolution of the DEM used for modeling. Using a higher resolution DEM for hydrology and hydraulic calculations is shown to remove some of the discrepancy introduced by lower resolution DEM data during the delineation process. Removing the cumulative effect of resolution is found to increase the goodness of fit for floodplains, typically by around five percent.

Further study is needed to understand what model types and DEM resolutions create the most realistic flooding event when compared to known events. There is also a need for further study of flood model combinations, and how modeling is affected by variations in location and land cover. The ultimate goal of the study is to further the ability of automated floodplain models to recreate accurate floodplains. This study analyzes the process, focusing on the variable of DEM resolution but the complexity of automated floodplain modeling requires additional future analyses.

Bibliography

Associated Press (AP)
    2008, Bush to Travel to Iowa to Inspect Midwest Flooding Damage, Associated Press, June 17, 2008

....
    2009, Obama Vows Help for Flood Victims, Associated Press, March 27, 2009

Aycock, William, Yong Wang
    2004, Comparison of the New Digital Flood Insurance Rate Map (DFIRM) with the Existing FIRM, Wilson, North Carolina. *Southeastern Geographer*, 44(2):159-169

Bates, P.D., A.P.J. DeRoo
    2000, A Simple Raster-Based Model for Flood Inundation Simulation. *Journal of Hydrology*, 236: 54-77

Berthier, E., T. Toutin
    2007, SPOT-HRS digital elevation models and the monitoring of glacier elevation changes in North-West Canada and South East Alaska. *Remote Sensing of Environment*, 112(5): 2443-2454

Callow, John, Kimberly Van Niel, Guy Boggs
    2006, How Does Modifying a DEM to Reflect Known Hydrology Affect Subsequent Terrain Analysis? *Journal of Hydrology*, 332(2007): 30-39

Deng, Y., J.P. Wilson, B.O. Bauer
    2007, DEM Resolution, Dependencies of Terrain Across a Landscape. *International Journal of Geographical Information Sciences* 21(2): 187-213

Federal Emergency Management Agency
    2005, *Map Modernization Guidelines and Specifications for Flood Hazard Mapping Partners*, www.fema.gov

Frech, Marshall
    2006, The Flood Safety Education Project. Retrieved April 8, 2008, from Flood Safety Web site: www.floodsafety.com Boulder CO

Frengs, M., N. Gruber, J. Sparks
    1999, *FEMA's Map Modernization and Revision of the Flood Insurance Study Process*, paper presented at the 23rd Annual Meeting, Association of State Flood Plain Managers, Portland OR

Gallant, John, Michael Hutchinson
	1997, Scale Dependence in Terrain Analysis, *Mathematics and Computers in Simulation*, 43: 313-321

Guisan, A, W. Thullier, N. Gotelli
	2005, Predicting species distribution: offering more than simple habitat Models, *Ecology Letters* 8(9): 993-1009

Hardy, R, P. Bates, M. Anderson
	1999, The Importance of Spatial Resolution in Hydraulic Models for Floodplain Environments, *Journal of Hydrology*, 216(1999): 124-136

Hawkins, B.A., D. Montoya, M.A. Rodriguez, M.A. Olalla-Tarraga, M.A. Zavala
	2007, Global models for predicting woody plant richness from climate: comment, *Ecology* 88(1): 255-259

Holmes, K, O. Chadwick, P. Kyriakidis
	2000, Error in a USGS 30-Meter Digital Elevation Model and its Impact on Terrain Modeling, *Journal of Hydrology*, 233(2000): 154-173

Horrit, M, P. Bates
	2001, Effects of Spatial Resolution on a Raster Based Model of Flood Flow, *Journal of Hydrology*, 253(2001): 239-249

Koster, Chris
	2008, Floods of '08, Retrieved September 2, 2008 from Missouri Attorney General Website http://ago.mo.gov/Environmental-Law-Blog/28/Floods_of_08/

Larson, Lee W.
	1997, The Great USA Flood of 1993; *Destructive Water: Water Caused Natural Disasters and their Abatement and Control*; IAHS 239(June 1996): 13-20

Li, Yingkui, Jacob Napieralski, Jon Harbor
	2008, A Revised Automated Proximity and Conformity Analysis Method to Compare Predicted and Observed Spatial Boundaries of Geologic Phenomena *Computers & Geosciences* December 2008 34:12 1806-1814

Madsen H.
	2000, Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology,* 235: 276-288

McMaster, Kevin
    2002, Effects of Digital Elevation Model Resolution on Derived Stream
    Network Positions, *Water Resources Research*, 38(4): 13.1-13.9

Missouri Geographic Information Systems Advisory Committee
    2008, *Does LiDAR Satisfy Your Data Requirements and Budget
    Constraints?* LiDAR White Paper #1, LiDAR workshop presentation June
    2008 Kansas City, Missouri

National Land Cover Dataset (NLCD)
    2001, Multi-Resolution Land Characteristics Consortium (MRLC), USGS,
    Retrieved Februrary 2, 2009 from http://www.mrlc.gov/nlcd.php

Napieralski, Jacob, Yingkui Li, Jon Harbor
    2006, Comparing Predicted and Observed spatial Boundaries of
    Geological Phenomena: Automated Proximity and Conformity Analysis
    (APCA) Applied to Ice Sheet Reconstructions *Computers & Geosciences*
    2006 32 124-134

National Academy of Sciences
    2009, More Accurate FEMA Flood Maps Could Help Avoid Significant
    Damages and Losses. *ScienceDaily* Retrieved April 10, 2009 from
    www.sciencedailly.com/releases/2009/01/090123111515.htm

National Digital Elevation Program
    2004, Guidelines for Digital Elevation Data Version 1.0, USGS, May 10,
    2004 Retrieved October 28, 2009 from
    http://www.ndep.gov/NDEP_Elevation_Guidelines_Ver1_10May2004.pdf

Noman, Nawajish, James Nelson, Alan Zundel
    2001, Review of Automated Floodplain Delineation From Digital Terrain
        Models, *Journal of Water Resources Planning and Management*, Nov-
        Dec: 394-402
....
    2003, Improved Process for Floodplain Delineation from Digital Terrain
        Models, *Journal of Water Resources Planning and Management*, Sept-
        Oct: 427-436

Omer, Creighton, James Nelson, Alan Zundel
    2003, Impact of Varied Data Resolution on Hydraulic Modeling and
    Floodplain Delineation, *Journal of the American Water Resources
    Association*, 39(2): 463-475

Pappenberger, Florian, K. Beven, K. Frodsham, R. Romanowicz, P. Matgen
2006, Grasping the Unavoidable Subjectivity in Calibration of Flood Inundation Models: A Vulnerability Weighted Approach, *Journal of Hydrology*, 333: 275-287

Peralvo, Manuel, David Maidment
2004, Influence of DEM Interpolation Methods in Drainage Analysis *GIS in Water Resources* CE 394K

Pielke Jr., R.A., M.W. Downton, J.Z. Barnard Miller,
2002, *Flood Damage in the United States, 1926-2000: A Reanalysis of National Weather Service Estimates*, Boulder, CO: UCAR

Shivakoti, Binaya Raj, S. Fujii, S. K. Boontanon, H. Ihara, M. Moriya, S. Tanaka
2008, Grid Size Effects on a Distributed Water Quantity-Quality Model in a Hilly Watershed, *Water Science & Technology*, 58.9: 1829-1836

Simon, Paul, Ryan Stack
2009, *Development of Dam Breach Innundation Studies Using Advanced Remote Sensing (LiDAR) for Emergency Action Planning*, Missouri Department of Natural Resources, Presentation at MMGUG meeting May 13, 2009, Jefferson City, MO

Stolz, A., C. Huggel
2008, Debris flows in the Swiss National Park: the influence of different flow models and varying DEM grid size on modeling results. *Landslides,* 5(3): 311-319

Szekely, B., D. Karatson
2004, DEM-based morphometry as a tool for reconstructing primary volcanic landforms: examples from the Borzsony Mountains, Hungary. *Geomorphology*, 63: 25-37

Tate, Eric, David Maidment, Francisco Olivera, David Anderson
2002, Creating a Terrain Model for Floodplain Mapping, *Journal of Hydrologic Engineering*, Mar-Apr: 100-108

TIME.com.
2008, Fifty Top 10 Lists of 2007: Top 10 Natural Disasters. Retrieved March 8, 2009, from Time Web site:
www.time.com/time/specials/2007/top10

Tribe A.
1992, Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method. *Journal of Hydrology,* 139: 263-293

United States Army Corps of Engineers
  2002, Hydrologic Engineering Center, *HEC-RAS Hydraulics Reference Manual*, www.hec.usace.army.mil, V.3.1

United States Geological Survey (USGS)
  2001, The National Flood Frequency Program - Methods for Estimating Flood Magnitude and Frequency in Rural and Urban Areas in Missouri. USGS Fact Sheet 015-01. http://pubs.usgs.gov/fs/fs-015-01/

Vazquez, R, J. Feyen
  2006, Assessment of the Effects of DEM Gridding on the Predictions of Basin Runoff using MIKE SHE and a Modeling Resolution of 600m, *Journal of Hydrology*, 334: 73-87

Walker, J, G. Willgoose
  1999, On the Effect of Digital Elevation Model Accuracy on Hydrology and Geomorphology, *Water Resources Research*, 35: 2259-2268

Wang, Yong, Tao Zheng
  2005, Comparison of Light Detection and Ranging and National Elevation Dataset Digital Elevation Model on Floodplains of North Carolina, *Natural Hazards Review*, Feb: 34-40

Wheater, H
  2002, Progress in and Prospects for Fluvial Flood Modeling, *Philosophical Transactions of the Royal Society of London*, 360: 1409-1431

Yang, Jie, Ronald Townsend, Bahram Daneshfar
  2006, Applying the HEC-RAS model and GIS Techniques in River Network Floodplain Delineation, *Canadian Journal of Civil Engineering*, 33: 19-28

Yilmaz, Musa, Nurunnisa Usul, Zuhal Akyurek
  2004, *Modeling the Propagation of DEM Uncertainty in Flood Inundation*, Paper presented at the 2004 ESRI conference