


November 2015

Using Digital Elevation Models Derived from Airborne LiDAR and Other Remote Sensing Data to Model Channel Networks and Estimate Fluvial Geomorphological Metrics

Noah Slovin
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/masters_theses_2

 Part of the [Fresh Water Studies Commons](#), [Geomorphology Commons](#), [Hydrology Commons](#), and the [Water Resource Management Commons](#)

Recommended Citation

Slovin, Noah, "Using Digital Elevation Models Derived from Airborne LiDAR and Other Remote Sensing Data to Model Channel Networks and Estimate Fluvial Geomorphological Metrics" (2015). *Masters Theses*. 297.

<https://doi.org/10.7275/7373499> https://scholarworks.umass.edu/masters_theses_2/297

This Open Access Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Using Digital Elevation Models Derived from Airborne LiDAR and other Remote Sensing
Data to Model Channel Networks and Estimate Fluvial Geomorphological Metrics

A Thesis Presented

by

NOAH B. SLOVIN

Submitted to the Graduate School of the University of Massachusetts Amherst in partial
fulfillment of the requirements for the degree of

MASTER OF SCIENCE

September 2015

DEPARTMENT OF GEOSCIENCES

Using Digital Elevation Models Derived from Airborne LiDAR and other Remote Sensing
Data to Model Channel Networks and Estimate Fluvial Geomorphological Metrics

A Thesis Presented

by

NOAH B. SLOVIN

Approved as to style and content by:

Christine Hatch, Chair

David F. Boutt, Member

Jon D. Woodruff, Member

Noah P. Snyder, Member

Julie Brigham-Grette, Department Head

Department of Geosciences

ABSTRACT

USING DIGITAL ELEVATION MODELS DERIVED FROM AIRBORNE LIDAR AND OTHER REMOTE SENSING DATA TO MODEL CHANNEL NETWORKS AND ESTIMATE FLUVIAL GEOMORPHOLOGICAL METRICS

SEPTEMBER 2015

NOAH SLOVIN, M.S. UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Dr. Christine Hatch

Recent advances in remote-sensing technologies and analysis methods, specifically airborne-LiDAR elevation data and corresponding geographical information system (GIS) tools, present new opportunities for automated and rapid fluvial geomorphic (FGM) assessments that can cover entire watersheds. In this thesis, semi-automated GIS tools are used to extract channel centerlines and bankfull width values from digital elevation models (DEM) for five New England watersheds. For each study site, four centerlines are mapped. *LiDAR* and *NED* lines are delineated using ArcGIS spatial analyst tools with high-resolution (1-m to 2-m) LiDAR DEMs or USGS National Elevation Dataset (NED) DEMs, respectively. *Resampled LiDAR* decreases LiDAR DEM resolution and then runs spatial analyst tools. National Hydrography Dataset (NHD) lines are mapped by the USGS. All mapped lines are compared to centerlines delineated from photography and LiDAR DEMs. Bankfull widths at each site are determined through three methods. Regional regression equations are applied using variables derived from LiDAR and NED DEMs separately, producing two sets of width results. Additionally, the *Hydrogeomorphological Geoprocessing Toolset* (HGM) is used to extract widths from LiDAR data. Widths are also estimated visually from aerial photos and LiDAR DEMs. Widths measured directly in the field or derived from field-data are used as a baseline for comparison.

I find that with a minimal amount of preprocessing, specifically through DEM resampling, LiDAR data can be used to model a channel that is highly correlated with the shape and location of the mapped channel. NED-derived channels model the mapped channel shape with even greater accuracy, and model the channel location only minimally less accurately. No tool used in this study accurately extracted bankfull width values, but analysis of LiDAR data by the HGM toolset did capture details that could not be resolved using regression equations. Overall, I conclude that automated, computerized LiDAR interpretation needs to improve significantly for the expense of data collection to be cost-effective at a watershed scale.

CONTENTS

	Page
ABSTRACT	iii
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
1 INTRODUCTION	1
1.1 Significance of Research	1
1.2 Background.....	2
1.2.1 Fluvial Geomorphologic Theory.....	2
1.2.2 Remote Sensing	6
1.2.3 Benefits of Remote Sensing to River Science and Management.....	9
1.2.4 Potential Use of Automated Extraction.....	9
1.3 Research Statement	10
1.4 Research Progression	10
2 STUDY AREAS	12
2.1 Watersheds.....	12
2.1.1 The Deerfield River Basin	12
2.1.2 The Fenton River Watershed.....	15
2.2 Data Sources	15
2.2.1 Ground-Truthed Data	15
2.2.2 USGS Data.....	16
2.2.3 LiDAR Data.....	17
2.2.4 Aerial Orthoimagery	18
3 METHODS.....	20
3.1 Automated Analysis	20
3.1.1 Introduction.....	20
3.1.2 Flow Routing and Accumulation.....	20
3.1.3 LiDAR Resampling.....	24
3.1.4 Regional Regression Equations.....	26
3.1.5 HydroGeoMorphological Geoprocessing Toolset.....	28
3.2 Manual Analysis.....	31
3.3 Overlaying Field Data	32
3.4 Analysis Products	34

3.4.1	Sinuosity Index.....	34
3.4.2	Centerline Offset	35
3.4.3	Width	36
4	RESULTS	37
4.1	Sinuosity Index Error	37
4.1.1	Introduction.....	37
4.1.2	Results	40
4.2	Centerline Offset Error	43
4.2.1	Introduction.....	43
4.2.2	Results	45
4.3	Bankfull Width Error.....	47
4.3.1	Introduction.....	47
4.3.2	Results	48
5	DISCUSSION.....	52
5.1	Centerline Mapping.....	52
5.1.1	Centerline Discussion	52
5.1.2	Centerline Modeling Sources of Error	53
5.2	Bankfull Width	55
5.2.1	Bankfull Width Discussion	55
5.2.2	Bankfull Width Sources of Error	57
6	CONCLUSIONS	59
6.1	Feature Extraction.....	59
6.1.1	Channel Planform	59
6.1.2	Bankfull Width	59
6.2	Further Work	60
6.2.1	Alternative Approaches	60
6.2.2	Tool Refinement	61
6.3	General Conclusions	63
	WORKS CITED.....	64

LIST OF TABLES

Table:	PAGE
2-1: LiDAR collection flight parameters.....	17
3-1: HGM Tool Parameters.....	31
4-1: Sinuosity Index Results Summary	40
4-2: Offset Error Results Summary.....	45
4-3: Bankfull Width Results.....	50

LIST OF FIGURES

Figure:	PAGE
1-1: Physical Characteristics of a River	2
1-2: "The Lane Diagram"	5
2-1: Map of Study Watersheds	12
2-2: Deerfield Basin Sub-Watershed Study Sites.....	13
2-3: Fenton River Study Site.....	15
3-1: Flow Routing and Accumulation	21
3-3: One-Meter DEM without Breaching	23
3-3: Breached One-Meter DEM	23
3-4: Breach, and Flow Routing and Accumulation Work Flow.....	23
3-5: LiDAR Resampling, Routing, and Accumulation Work Flow	25
3-6: Regional Curve ArcGIS Workflow	28
3-7: Binary River Output	29
3-8: River Width Tool Schematic	29
3-9: Overlaying Field Data	32
3-10: Manual Width Measurement.....	33
3-11: Sinuosity Index Calculation.....	34
3-12: Centerline Offset Calculation	35
4-1: Sinuosity Index Error Distribution	37
4-2: Sinuosity Index Error Summary	38
4-3: Chart of Modeled Sinuosity over Measured Sinuosity.....	39
4-5: Deficient Sinuosity in a Sinuous Reach	42
4-4: Excess Sinuosity in a Straight Reach	42
4-6: Normalizing Centerline Offset.....	43
4-7: Percent-Bank-Distance Error Distribution	44
4-8: Offset-Error Summary	45
4-9: Erroneous Model within Banks	46
4-10: Bankfull Width Error Distribution.....	48
4-11: Example of Varying HGM Tool Results, Fenton Brook	49
4-12: Bankfull Width Results Summary.....	50
5-1: Model Failure	53
5-2: Dataset Inconsistency	54
5-3: Visual Estimate Results.....	55
5-4: Regional-Curve Results: Modeled Width over Measured Width.....	56
5-5: HGM Width Results: Modeled Width over Measured Width.....	57

CHAPTER 1

INTRODUCTION

1.1 Significance of Research

Rivers are shaped by complex combinations of forces dependent on climate and geology, such as patterns in precipitation, topography, land-cover, and soil types. The interactions of these many parameters create the processes of erosion and deposition that determine a stream's morphology. Fluvial geomorphology (FGM) is the study of those processes and the landforms they create. A thorough understanding of FGM, geology, and climate can aid in the understanding of sediment transport patterns, erosion rates, and aggradation and degradation processes (Rosgen, 1994). Having a grasp of these processes in turn gives important insight to allow for the prediction of river responses to changes in sediment supply and discharge (Montgomery & Buffington, 1998; Rosgen, 1994).

Incorporating this insight into river management, restoration and mitigation projects, land-use planning, and policy can prevent costly damage, increase sustainability and resiliency, and avoid unintended impacts up- or downstream. Additionally, the concepts of FGM present a method of river classification based on quantifiable physical attributes. This is important for comparison between similar types of streams, and for determination of a stream's "state" relative to a pristine reference condition. Having a common language for describing stream type and condition is essential for establishing large-scale frameworks for the protection of water resources, improvement of aquatic habitat, and the monitoring and administration of such programs (European Commission, 2000). It also creates a mechanism for prediction of stream evolution, such that one can predict that a certain type

of stream will, given some change in parameters, evolve into another stream type (Brierley, Fryirs, & Cohen, 1996; Raven, et al., 2002; Rosgen, 1994).

1.2 Background

1.2.1 Fluvial Geomorphologic Theory

The physical, chemical, and biological processes, reactions, and forces that interact with one-another within a river system are hugely complex. Yet one of the central premises of FGM is the simple idea that flowing water moves sediment, and changes in the flow of water will change the way sediment is eroded, transported, and deposited. The physical characteristics of a stream are determined by geology, mineralogy, biology, climate, and more, but ultimately they are inextricably linked to the flow of water through the system. Changes in the stream's flow patterns along its course affect the channel's physical characteristics, and those physical characteristics, in turn, affect the flow patterns of water and sediment through the system (Montgomery & Buffington, 1998; Rosgen, 1994; Rosgen, 1996). Thus, characterizing the processes of a river provides insight as to the

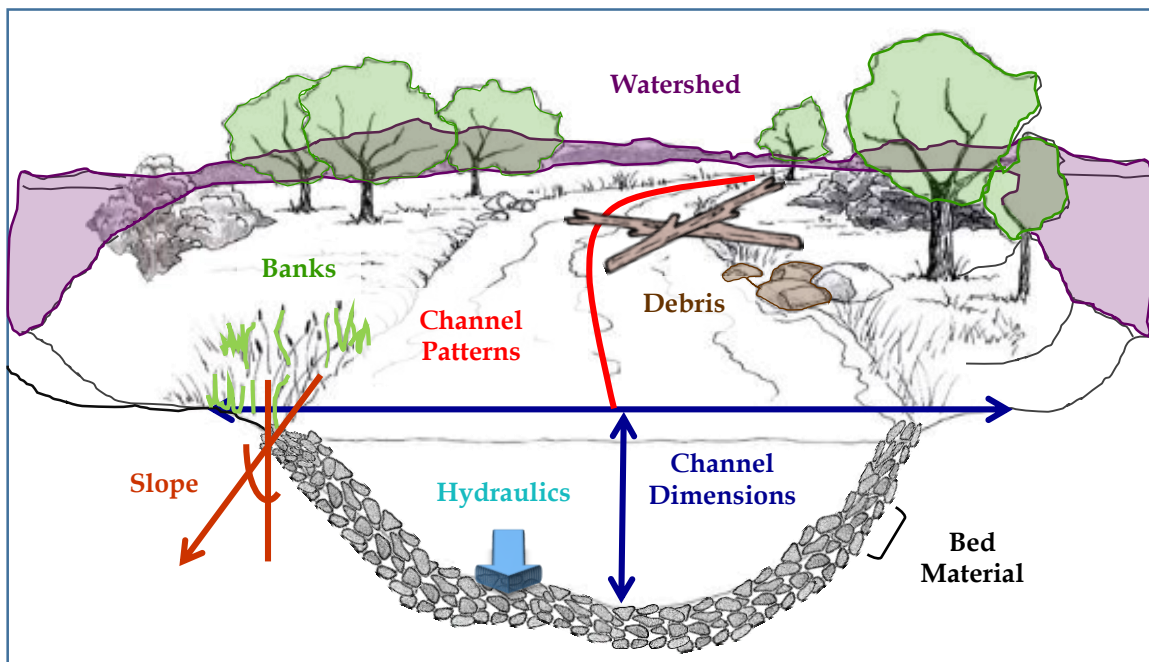


Figure 1-1: Physical Characteristics of a River

Image Altered From TERC: http://www.concord.org/~btinker/GL/web/water/rivers_streams.html

form of the river, and characterizing the form of a river provides insight into its processes.

At a given point in a stream system, the water flowing across that site has a certain amount of energy that can be used to do work on the bed and banks of the channel. This energy is divided between the water's potential energy due to its elevation (PE_z), its potential pressure energy created by the accumulation of water into a mass with depth h (PE_p), and its kinetic energy (E_k). Other forms of energy contained within the water column, such as temperature or chemical energy, are not considered here because their contributions to sediment erosion and transport are insignificant (Montgomery & Buffington, 1998). The equation describing the available energy of the water column is:

$$\text{Equation 1} \quad E_w = PE_z + PE_p + E_k = (m \times g \times z) + (m \times g \times h) + \frac{(m \times u^2)}{2}$$

where E_w is the total energy of the water at that site, m is the mass of the water in the column, g is the acceleration due to gravity, z is the land surface elevation, h is the depth of the water column, and u is the average water velocity. As water flows downstream, some of this energy is lost as it is used to do the work of entraining and transporting channel material. Over a short stream length ΔL , where the change in water velocity can be assumed to equal zero ($\Delta u = 0$) and gravity and the water column mass can be assumed to remain constant, the loss of energy can be written as:

$$\text{Equation 2} \quad \frac{\Delta E_w}{\Delta L} = \frac{\Delta(m \times g \times z) + \Delta(m \times g \times h) + 0}{\Delta L} = (m \times g) \frac{\Delta(z+h)}{\Delta L}$$

Note that $(z + h)$ is the ground surface elevation plus the water depth, or simply the elevation of the water surface. Therefore $\frac{\Delta(z+h)}{\Delta L}$ is the slope S of the water surface, and the equation can be rewritten as:

Equation 3
$$\frac{\Delta E_w}{\Delta L} = m \times g \times S$$

The mass of the water column is equal to the product of water density and the column volume, which when investigating a channel cross section is equal to the cross-sectional area multiplied by one unit-length, giving:

Equation 4
$$\frac{\Delta E_w}{\Delta L} = \rho_w \times A \times \Delta L \times g \times S$$

This energy available to do work is balanced by the shear resistance of the channel banks and bed:

Equation 5
$$\tau_b = \tau_0 \times \Delta L \times P$$

where τ_b is the shear resistance of the bed and banks, τ_0 is the total shear stress that the water exerts on the bed and banks, and P is the wetted perimeter of the cross section.

Writing out the balance of forces gives:

Equation 6
$$\tau_0 \times \Delta L \times P = \rho_w \times A \times \Delta L \times g \times S$$

which can be solved for the shear stress to give:

Equation 7
$$\tau_0 = \rho_w \times \frac{A}{P} \times g \times S$$

Equation 7 describes the force that the water flowing across a unit length of stream channel exerts on the beds and banks of a channel. This force drives the erosion and transport of sediment throughout a stream system (Montgomery & Buffington, 1998).

Bagnold (1966) previously developed a similar equation, but included the discharge Q of

the stream. This equation defines the parameter of stream power as:

$$\text{Equation 8} \quad \Omega = \rho_w \times g \times Q \times S$$

Where stream power Ω , or the rate at which energy is expended on the stream banks and bed, is defined as a product of water density ρ_w , gravity g , discharge Q and channel slope S .

Discharge itself is based on the amount of water that is flowing to a particular point from upstream:

$$\text{Equation 9} \quad Q = k \times A^c \quad (\text{Dunne \& Leopold, 1978})$$

Here, A is the upstream area, k is the depth of precipitation at each unit area, and c is an empirically derived dependency value that varies by scale, lithology, land use, and climate, and which accounts for precipitation that may not reach the stream channel.

The relationship between discharge, slope, and sediment flow is also noted in Lane's landmark 1955 paper, where he conceptualized how sediment flow and sediment size are balanced by water flow and stream slope (Lane, 1954):

$$\text{Equation 10} \quad Q_s \times d \propto Q_w \times S$$

Where Q_s and Q_w are sediment flow and water flow, respectively, d is sediment clast diameter, and S is stream slope.

Based on these equations, it is clear that channel slope, discharge, and cross-sectional

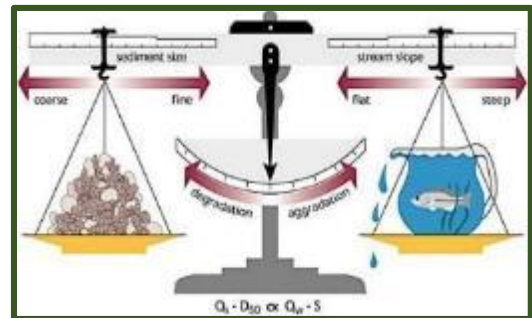


Figure 1-2: "The Lane Diagram"
Depicts balance between sediment flow and water energy.
From Rosgen (1996) based on Lane, 1955).

geometry are key parameters for determining stream power and therefore for predicting stream behavior across a landscape. For my thesis, I focused on channel slope S and channel bankfull width W (as a single, important, and accessible aspect of cross-sectional geometry). Bankfull width is the width of a flowing channel during bankfull flow, defined as "...the discharge at which moving sediment, forming or removing bars, forming or changing bends and meanders, and generally doing work results in the average morphologic characteristics of channels." (Dunne & Leopold, 1978). Its importance to FGM characterization of a river is evident in its use throughout the literature and in existing physical stream assessment methods (Brierley, Fryirs, & Cohen, 1996; Committee on Floodplain Mapping Technologies, 2007; Environment Agency, 2003; Montgomery & Buffington, 1998; Rosgen, 1994; Rosgen, 1996).

1.2.2 Remote Sensing

Remote sensing, generally, is the collection of information about any target without directly coming into contact with that target. In practice, the phrase describes data collection via sensors that detect and evaluate electromagnetic energy. There are two overarching categories of remote-sensing methods: passive remote sensing, where the sensor evaluates radiation emitted from an external source; and active remote sensing, in which the sensor platform emits radiation itself. Data can be collected through sensor platforms mounted on satellites, aircraft, watercraft, vehicles, or tripods, or simply carried (Carbonneau & Piégay, 2012).

In addition to there being a variety of tools and methods for remote data-collection, there is a variety of types of data that one might collect. Data that can be used to create a digital elevation model of a landscape allows a researcher to model the fluvial processes occurring in that landscape, as opposed to imagery data, for example, where form can be

observed but not necessarily processes. In my thesis I utilized the strengths of digital elevation data. Following is a summary of common methods of collecting elevation data.

Photogrammetry is the use of aerial photography to produce both planimetric and three-dimensional elevation maps (Committee on Floodplain Mapping Technologies, 2007). This method generally limits light detection to visible wavelengths, though infrared and thermal radiation sensors are also used in airborne- or satellite-based passive remote sensing (Carbonneau & Piégay, 2012). Images collected by aircraft or satellites are rectified to remove tilt and relief displacement, and then heights are calculated from correlating overlapping images. (Committee on Floodplain Mapping Technologies, 2007).

Radar and LiDAR systems are active sensing methods. Radar units operate in the one to ten centimeter wavelength range, while LiDAR units typically emit light at 1024 nm. These signals are reflected off of surfaces back to the sensor, and those return signals are analyzed (Committee on Floodplain Mapping Technologies, 2007). LiDAR will be covered in more detail below.

Interferometric Synthetic Aperture Radar measures the intensities and round-trip times of microwave signals (3-40,000 MHz) reflecting off of a surface. The interferometry aspect of this method indicates the use of two emission and detection apertures, allowing for an enhanced three-dimensional view. The synthetic aperture refers to the aspect of this method that synthesizes a very long antenna, improving the resolution. Derived from topographic data collected in the field. Most of this data has a 30 m to 10 m horizontal resolution. Vertical resolution is very variable (Carbonneau & Piégay, 2012).

1.2.2.1 *LiDAR*

LiDAR is an important and relatively new technology being used for fluvial geomorphic assessment. A LiDAR unit determines a surface's topography by emitting laser pulses and

measuring the time it takes for the laser to bounce back. LiDAR units can be used on the ground or on a boat, but for a spatially large-scale assessment, a unit can be mounted on a low-flying aircraft and flown over the area of interest. This aerial laser scanning method can have a vertical accuracy of 5 to 30 cm with a horizontal resolution of 20 to 80 cm (Hohenthal, Alho, Hyypä, & Hyypä, 2011; Carbonneau & Piégay, 2012).

As an additional benefit, the LiDAR unit sends out multiple laser pulses that each have a small areal footprint. Some of this light reflects off of treetops, airborne objects, and other obstacles, while some passes through these barriers and reflects off the ground surface, allowing LiDAR systems to “see” through trees (Hohenthal, Alho, Hyypä, & Hyypä, 2011). Signal processing can be used to analyze “bare earth” data independently. This high-resolution, rapid data collection method can be and has been used in a variety of geomorphic studies. LiDAR is used to study landslides, recognize fluvial depositional features, and study stream longitudinal profiles (Cavalli, Tarolli, Marchi, & Fontana, 2008; Sinha, 2000).

In the context of FGM, LiDAR can be used to measure numerous important parameters, such as reach length, sinuosity, connectivity, terrace elevations, bank incision, water surface elevation, valley cross-sections, landslides, road density, stream crossings, impervious surfaces, and vegetation height. To a limited degree, LiDAR can also be used to determine channel cross-sections, bankfull depth, bank angle, bank type, bank stability, pool frequency and length, and large woody debris within the channel (Faux, Buffington, Whitley, Lanigan, & Roper, 2009; Kasprak, Magilligan, Nislow, & Snyder, 2012). LiDAR is especially useful for mapping land cover, an important part of the hydrological systems controlling stream morphologies (Snyder, 2009).

A number of high-resolution elevation datasets of New England region watersheds have

been collected within the last decade using airborne LiDAR. This data can give elevation values that are an order-of-magnitude more finely-scaled, both vertically and in the planform, than previously available data. I used this dataset in my research to compare its outputs with those of more coarse data.

1.2.3 Benefits of Remote Sensing to River Science and Management

The complexity and interdependency of an entire stream's fluvial geomorphic character necessitates an assessment of a watershed as a whole in order to fully model the past, present, and future processes occurring at any one point. Traditional assessment methods utilize a combination of general data collection from remote-sensing technologies, and detailed data gathering in the field (Brierley, Fryirs, & Cohen, 1996; Environment Agency, 2003; Montgomery & Buffington, 1998; Rinaldi, Surian, Comiti, & Bussettini, 2013; Rosgen, 1994). These are time-consuming, expensive, often subjective methods, and impossible to complete over a large watershed on a useful time-scale without a massive staff and budget.

1.2.4 Potential Use of Automated Extraction

Improvements in both the technology used to collect data remotely and the processing used to interpret that data present an opportunity for researchers to assess entire watersheds at high resolutions, quickly and relatively inexpensively. Much of this work still requires subjective interpretations of data, such as aerial photographs, performed by individual researchers. An alternative is to use basic fluvial-geomorphic principles, in conjunction with relatively unchanging data such as topography, to objectively calculate consistent metric values. Again, decreasing the time and cost associated with a watershed assessment, while increasing the consistency and repeatability of results, is essential to the successful management of that watershed. The use of automated interpretation

programs applied to digital elevation models and is the subject of my thesis.

1.3 Research Statement

Identification of stream processes, including those related to habitat, water quality, and flood hazards, is important. Scientists and engineers are able to predict patterns of behavior and changes within stream systems by accurately measuring key physical, fluvial-geomorphic parameters. Because of the necessarily interconnected nature of stream systems, it is often useful to map these parameters across an entire watershed in order to more completely understand the processes as work and more effectively prioritize management efforts.

Remotely-sensed elevation data can be used to determine fluvial geomorphic information about a channel and its watershed to a degree of detail and accuracy that is useful for assessment and management purposes, and that this can allow for faster assessment of larger areas at lower cost. New technologies and data-analysis methods create an opportunity to measure important physical metrics accurately, consistently, and relatively inexpensively, over a large spatial scale.

High-resolution airborne-LiDAR elevation data and available GIS-based automated analysis methods allow for the collection of accurate and precise fluvial-geomorphic data at a watershed scale.

1.4 Research Progression

To assess my thesis statement, I used digital-elevation models of different resolutions, in combination with ArcGIS-based tools that are capable of automated performing relevant analyses, to derive a set of fluvial-geomorphometric parameters for a number of watersheds. In the following pages, I describe that process, and then compare the results

to data collected by professionals in the field. Based on these results, I determine the effectiveness and limitations of remote-sensing technologies and interpretation methods in order to add to the ongoing growth and improvement of the watershed assessment field.

My research consisted of the following steps:

1. Determine a small number of fluvial-geomorphic metrics that play a significant role in determining stream power, on which this thesis will focus.
2. Compile a comprehensive list of automated extraction tools for deriving those metrics from digital-elevation models.
3. Use a selection of those tools to interpret existing elevation data on watersheds for which both high-resolution LiDAR data and field data are available. Use lower-resolution National Elevation Dataset models in addition to LiDAR elevation data in extraction.
4. Compare results of the various methods and data sources to note differences based on the program and the data resolution.
5. Compare those results to data collected in the field or from other sources such as aerial photographs. Assess the effectiveness and limitations of these automated interpretation methods, taking into account availability of data and tools.

CHAPTER 2

STUDY AREAS

Five watersheds of varying sizes were chosen from around the New England area based on the availability of both LiDAR elevation data and field-based observations. Four of those are sub-watersheds of the Deerfield River basin on the western side of the Connecticut River in Massachusetts and southern Vermont. The last is the Fenton River, which neighbors the campus of the University of Connecticut in Storrs, Connecticut.

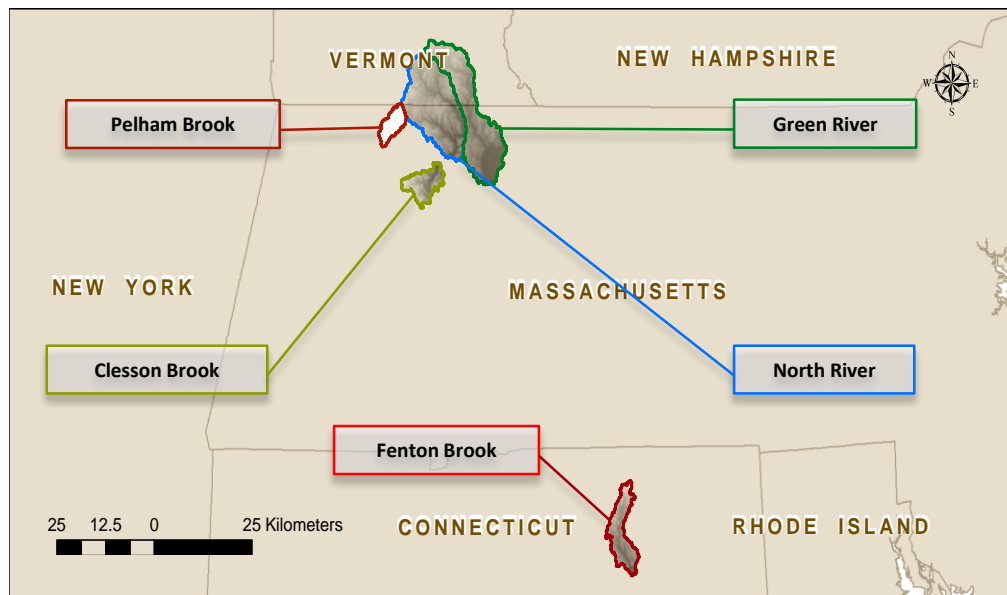


Figure 2-1: Map of Study Watersheds

2.1 Watersheds

2.1.1 The Deerfield River Basin

The Deerfield Basin brings water to the Connecticut River from the west, covering about 1,722 square kilometers in Massachusetts and southern Vermont. Nearly one-thousand kilometers of stream bring water from 866-meter high headwaters through steep narrow valleys in the uplands until a more gentle reach within the Connecticut River Valley, before joining that river. Along its course the river and its tributaries flow through 36 towns, and

on its main-stem through 19 flood-control and hydroelectric dams. Approximately 78 percent of the watershed is forested, three percent is urban, and the rest is agricultural (Deerfield River Watershed Association, 2014).

This river, its tributaries, and the land it drains, have long histories of human use and alteration, from deforestation and log-drives, to hydroelectric dams and channel straightening. Record-high precipitation from Hurricane Irene in 2011 highlighted risks associated with the river, including flooding, channel migration, road and bridge failures, and large landslide.

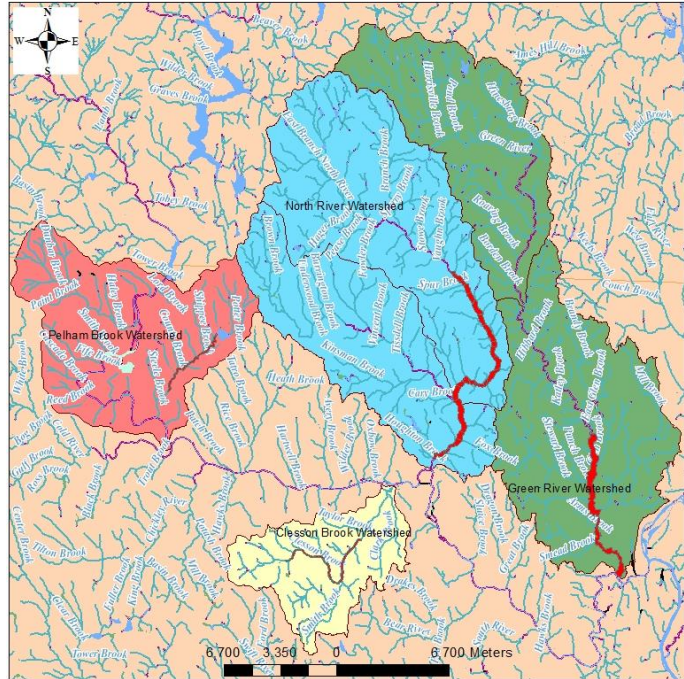


Figure 2-2: Deerfield Basin Sub-Watershed Study Sites

The sub-watersheds assessed in this project are the North River, Green River, Pelham Brook, and Clesson Brook.

2.1.1.1 *The North River*

The North River drains around 240 square kilometers in northern Massachusetts and southern Vermont. It is moderately steep, dominated by riffle-pool morphologies and cobble substrates, with bankfull widths ranging from 15 to 42 meters (McDonough, Mabee, & Marcus, 2013b). The river flows from Halifax and Whitingham to Shelburne, Massachusetts, through hilly land that is 83% forest, 9% agriculture, and 3% residential. (MEEA, 2004).

2.1.1.2 *The Green River*

The Green River watershed covers 230 square kilometers from Marlboro, Vermont, to Greenfield, Massachusetts. Land use in Massachusetts is 65% forest, 13% agriculture, and 11% residential, with most of the residential land in the downstream end of the river as it flows through Greenfield. This downstream, urbanized area is also where high concentrations of crossings, structures, and infrastructure can be found. The river channel cuts through floodplain alluvium, glacial lake sediment, and abandoned stream terraces, has a bed of mostly sand and gravel, and 16.5 to 41-meter bankfull width values (McDonough, Mabee, & Marcus, 2013c; MEEA, 2004).

2.1.1.3 *Pelham Brook*

Pelham Brook is a 35 square-kilometer watershed that drops through a steep, confined valley. The predominantly cobble channel is mostly made up of riffle-run morphologies and has a slope averaging 1.3% and bankfull widths ranging from 9 to 24 meters (McDonough, Mabee, & Marcus, 2013d). The watershed extends from Rowe to Charlemont, Massachusetts, and is 87.1% forest, 4.0% agriculture, 3.9% residential (MEEA, 2004).

2.1.1.4 *Clesson Brook*

Beginning in Hawley, Massachusetts, Clesson Brook runs down steep terrain (slope averaging 2.1%) to Buckland, Massachusetts, then a low-gradient (slope averaging 1.4%), wide floodplain area with high levels of agricultural activity until the joins the Deerfield, also in Buckland (MEEA, 2004). The 55 square-kilometer watershed is approximately 81.4% forest, 9.6% agriculture, and 4.7% open land, with the channel bed consisting mostly of cobble and gravel substrate, and a bankfull width ranging from 9 to more than 25 meters (McDonough, Mabee, & Marcus, 2013a; MEEA, 2004).

2.1.2 The Fenton River Watershed

Partially bordering the east side of the University of Connecticut in Storrs, Connecticut, the Fenton River watershed is a 90-square-kilometer sub-watershed of the Natchaug River, a tributary of the Thames River basin. It begins in State Forest land in Willington, Connecticut, and meets the Natchaug in Mansfield. Land use is 74% forest, 14% urban, and 8% agricultural (CTDEEP, 2012). Bankfull widths range

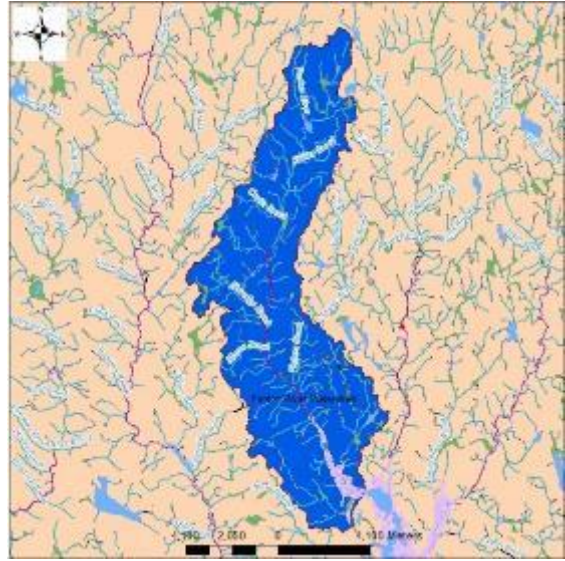


Figure 2-3: Fenton River Study Site

from around 7 to 27 meters (Brink, Skewes, & Henry, 2013; Lamont, Farrell, Walker, & Rosa, 2013; Pivarnik, Nicoulin, & MicCusker, 2013).

2.2 Data Sources

2.2.1 Ground-Truthed Data

New England Environmental Inc., University of Massachusetts Amherst, and the Massachusetts Geological Survey performed fluvial geomorphic assessments in order to map erosion hazards along these streams. Methods were derived from the U.S. EPA Rapid Bioassessment protocols, the Center for Watershed Protection's Unified Stream Assessment, Rosgen's assessment and classification methods, the Vermont Stream Geomorphic Assessment Protocols, and the Generic QAPP for Stream Morphology Data Collection. Field work was performed in the summer of 2012. (McDonough, Mabee, & Marcus, 2013a,b,c,d).

Graduate students at University of Connecticut, under the supervision of Dr. Will Ouimet, collected fluvial geomorphic data for 3 separate reaches of Fenton River as part of a geography course. Data collection and research covered historic aerial photographs and land uses, recent LiDAR data, cross-sections collected using total stations in the field, and field characterization of grain sizes, bars, and erosional and depositional features (Brink, Skewes, & Henry, 2013; Lamont, Farrell, Walker, & Rosa, 2013; Pivarnik, Nicoulin, & MicCusker, 2013). In the end only two of these datasets proved useful as baseline field comparisons, as I was unable to find the raw cross-sectional data that I required for the third.

2.2.2 USGS Data

In addition to the high-resolution field-data and LiDAR-based data, I included National Hydrography Dataset (NHD) and Watershed Boundary Dataset (WBD) shapefiles in my analysis, as well as terrain-models from the National Elevation Dataset (NED). I used the WBD polygons as bounding areas for my assessments, allowing me to analyze only the watershed of interest, rather than much larger elevation datasets. I included NHD channel centerlines in my analysis of channel centerlines so that I could compare the results of extracting data from high-resolution elevation models against the coarser-resolution but readily available NHD data. NHD data is provided in vector form at a nominal scale of 1:24,000, and can be downloaded from the USGS National Map Viewer at <http://nationalmap.gov/>.

I used 1/3 arc-second NED digital elevation models as the base-data for performing some of the same fluvial-geomorphic analyses as I performed on the LiDAR DEMs. As with the NHD vectors, I used this data to compare high-resolution DEMs to coarser-resolution but readily and widely available DEMs. For the New England areas I studied, the raster cell is

approximately 9.07 meters in both the x- and y-direction. The USGS NED is a constantly evolving dataset, with upgrades and improvement being made in a patchwork manner. The $\frac{1}{3}$ arc-second elevation data is among the oldest and poorest quality. The study watersheds I assessed in this thesis fall within fifteen different USGS 7.5-minute quadrangles. A contour-to-grid technology called LineTrace+ was used to digitize 1:24,000-scale cartographic contour maps for each quadrangle. These contour maps were in turn each produced separately, using different methods at different times. Drainage patterns were enforced by using NHD streamlines as elevation breaklines and water-body boundaries as contour lines (Osborn, et al., 2001). Contour maps were digitized in the 1990s, but the contour maps themselves range in production date from 1944 to 1981 (USGS). RMSE values for DEMs digitized in this manner have been found to be +/- 6.0 to +/- 7.0 meters (Chirico, 2004).

2.2.3 LiDAR Data

The LiDAR data I used came from two different LiDAR collection projects, each performed for different customers and by different companies, and each covering a different area.

Flight Parameter	D.H.H. Project	Northeast CT Project
Sensor	Optech ALTM 203	Leica ALS60 sn146
Laser Rate	50 KHz	118 KHz
Nominal Side Lap	30 %	50 %
Avg Point Distribution	2 points per meter	2 points per meter
Horizontal Resolution	2 meters	1 meter
Vertical Resolution	0.15-0.4 meters	0.185 meters

Table 2-1: LiDAR collection flight parameters

2.2.3.1 *Deerfield-Hudson-Hoosic Project*

In March and April of 2012, the Northrop Grumman Advanced GEOINT Solutions Operating Unit was used to collect LiDAR as part of the FEMA Hudson-Hoosic Deerfield LiDAR project. The study covers 2,895 square miles in northwestern Massachusetts, southwestern Vermont, and eastern New York. An Optech ALTM203 airborne LiDAR sensor was flown approximately 7,500 feet above ground level, collecting elevation data

at an average spacing of 2 points per square meter. After processing, bare-earth digital elevation data had a horizontal resolution of 2 square meters, and a vertical resolution between 0.15 and 0.4 meters (Northrop Grumman Corporation, 2012).

Data was downloaded from the Massachusetts Office of Geographic Information (MassGIS) and Vermont Center for Geographic Information (vcgi.vermont.gov).

2.2.3.2 *Northeast Connecticut Project*

The architecture and engineering consulting firm Dewberry managed a LiDAR collection and processing project for the USDA Natural Resource Conservation Service under a US Army Corps of Engineers contract, in 2010. Earth Eye flew a Leica ALS60 sn146 LiDAR unit in November and December of that year, collecting elevation data across 4,589 square kilometers of the Connecticut counties of Tolland, Windham, Hartford, Middlesex, and New London. LiDAR return density was approximately 2 points per square meter. The post-processing bare-earth digital elevation model had a horizontal resolution of 1 square meter and an average vertical resolution of 0.185 m (Dewberry, 2011).

Data was provided through personal communications with Dr. Will Ouimet (February, April, 2014).

2.2.4 Aerial Orthoimagery

As a further check on the accuracy of the automated tools I tested in this thesis, I visually compared results against aerial-photographs of the study areas. These orthorectified images came from different sources and dates, depending on the site. The Deerfield Watershed is covered by aerial photographs taken in March of 2012, with the exception of the Green River study area, which has imagery from April 2011. These images were provided with the other Deerfield Watershed data from the Massachusetts Geological Survey projects (McDonough, Mabee, & Marcus, 2013a; 2013b; 2013c; 2013d). Fenton

River aerial imagery came from the Connecticut Department of Energy and Environmental Protection's online data archive and was taken in 2012 (USGS, 2013).

CHAPTER 3

METHODS

3.1 Automated Analysis

3.1.1 Introduction

A geographic information system (GIS) allows for the type of spatial processing and analysis needed to understand rivers at the watershed scale. For this project I used ESRI ArcGIS (ESRI, 2014).

The LiDAR and NED DEMs covered areas much larger than each single watershed, and so my first step was to limit the extent of my ArcGIS processing to just those areas. I used the USGS National Hydrology Dataset Watershed Database (WDB) Hydrologic Unit Code 12 (HUC-12) watershed polygons to demarcate the general region of interest. To avoid any loss of data due to misrepresented watershed boundaries in the WDB, I expanded these polygons by one kilometer to create a mask, which I used to extract just the area of interest from the DEMs.

3.1.2 Flow Routing and Accumulation

There are a number of ways to identify stream channels on a DEM, but for the purposes of hydrologic analysis of a watershed, it is most effective to use the elevation data to route flow over the landscape. As mentioned earlier, this method models the processes that form the river, rather than modeling the form of the river and then backing out the processes.

Digitized terrain can be analyzed in order to determine the flow paths that water on the surface would take. From that information the accumulation of flow can also be calculated.

The cells in low valleys that are found to have high amounts of flow accumulation are designated as stream channels. The channel centerline drawn by this process can be used to assess the channel planform.

3.1.2.1 Basic Flow Routing and Accumulation in GIS

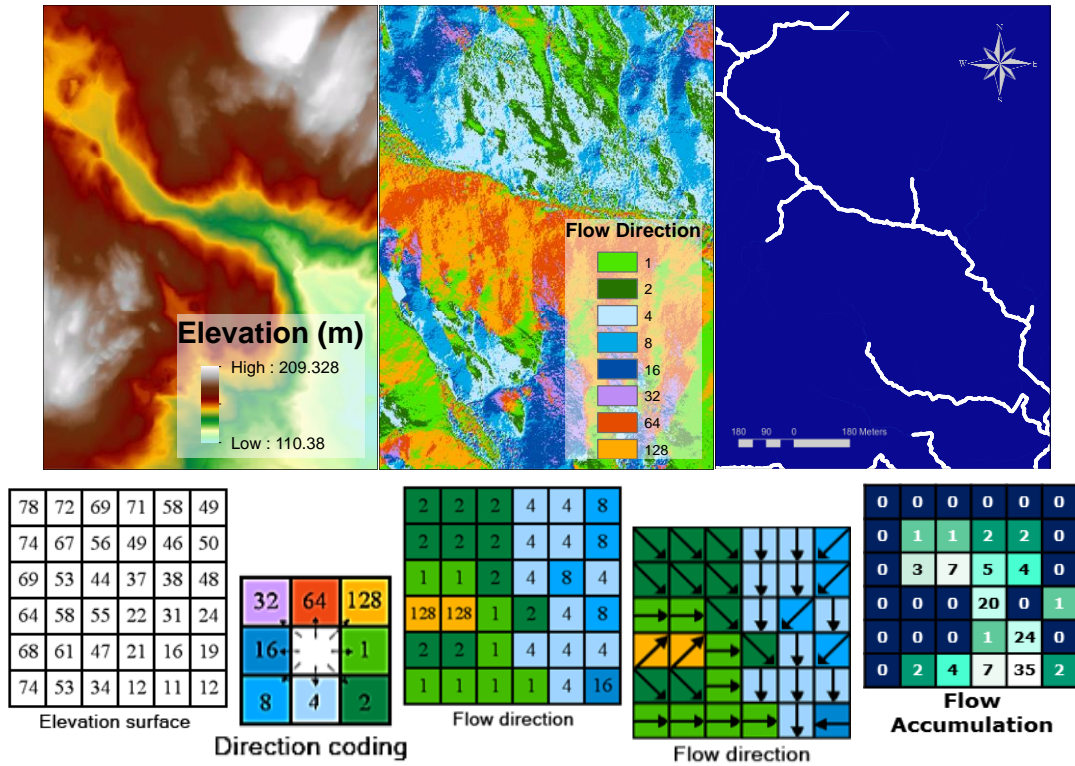


Figure 3-1: Flow Routing and Accumulation
 ArcGIS-images and schematic representations of the flow routing and accumulation process. Left-most figures represent a digital elevation model, center figures represent flow-direction determination, and right-most figures represent flow accumulation and channel delineation.

For this project the previously described flow routing and accumulation processes were performed with the ArcGIS Hydrology Toolset in spatial analyst. This process includes running the “Flow Direction” tool, which assigns each cell one of eight values, depending on which adjacent cell has the lowest elevation relative to that central cell (“D8 Algorithm” (O’Callaghan & Mark, 1984)). If the central cell is lower than all adjacent cells, though, there is no flow off that cell, and it becomes a “pit.” The “Fill” tool ensures hydrologic connectivity by raising the elevation of both erroneous and real “pit” pixels until they are at the same altitude of the pixel that had been blocking flow. This is especially helpful

when using a LiDAR dataset, which is liable to have a large amount of pits caused by noise in the data due to its high resolution and reliance on preprocessing filtering.

Basic flow routing, therefore, consists of using the “Fill” tool to remove pits, the “Flow Direction” tool to calculate the direction of flow that water on each cell would follow onto an adjacent tool, and the “Flow Accumulation” tool, which uses the flow direction raster created in the previous step to determine the amount of flow coming into each cell. Areas of high accumulation are determined to be stream channels by assigning a threshold value of accumulation. This last step is accomplished using the “Con” tool and re-assigning cells with an accumulation value below the user-defined threshold with a new value of zero. These processes allow one to draw a channel network across a landscape.

3.1.2.2 *Breaching Obstacles*

While many pits and obstacles in a LiDAR-based DEM come from noise and errors, there may also be real structures such as bridges and culverts that should not be blocking flow, but will result in a modeled environment similar to a dam if “Fill” is used. These obstacles must be removed from the dataset in order for the ArcGIS flow routing algorithm to create a flow-path that follows the channel.

I removed flow obstacles using an ArcGIS toolset called “ToBreachThroughBridges,” provided by Guenole Chone of Concordia University in Montreal (personal communication, May 12, 2014). To use this model, I first manually found all of the obstacles within the stream channel, and then drew polygons that covered those obstacles. The model creates a new DEM in which every cell covered by a polygon is assigned a new elevation equal to the lowest-elevation pixel covered by any polygon. The “Fill” tool is then able to raise the elevation of these pits until water can flow across them and continue downstream.

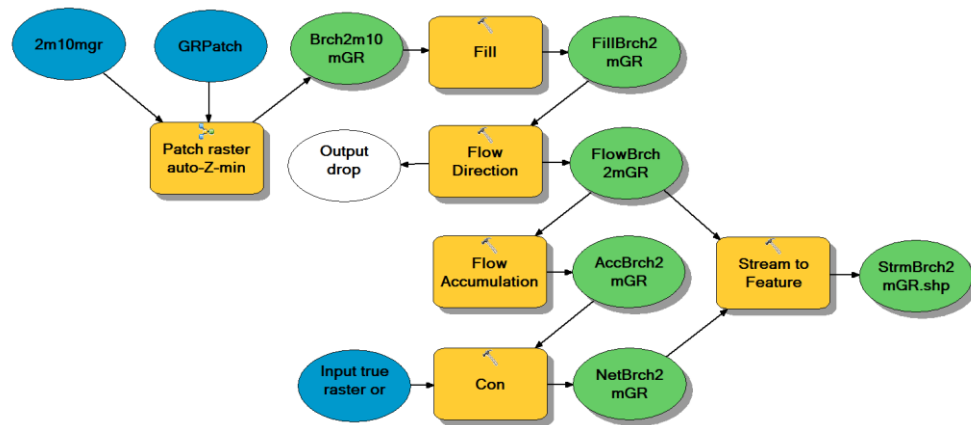


Figure 3-2: Breach, and Flow Routing and Accumulation Work Flow

When obstacles exist in a DEM, the “Fill” tool creates flat areas upstream of them. When the “Flow Direction” and “Accumulation” tools are run on such a dataset, the resulting channel lines are distinctly straight through these flat areas, making them easy to spot. Centerlines mapped parallel to one another is also indicative of flow obstacles, such as a road constructed alongside a channel. These two patterns make it easier to find flow obstacles than from direct inspection of a DEM, and so I ran these flow-routing processes on the raw DEM first, in order to find obstacles. I drew polygons over the obstacles so that the widths were close to or less than the channel bankfull (based on visual assessment), and their lengths extended above and below the obstacle so that the polygon covered some of the channel bottom.

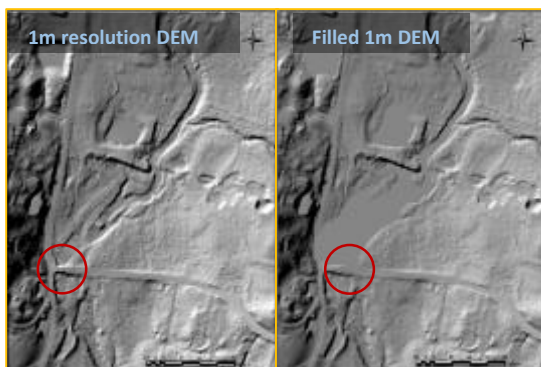


Figure 3-3: One-Meter DEM without Breaching

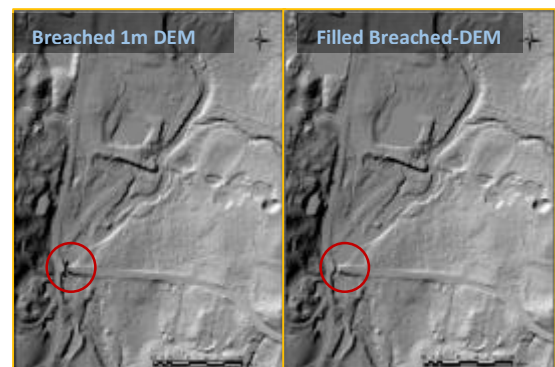


Figure 3-4: Breached One-Meter DEM

I ran the “ToBreachThroughBridges” model provided by Chone, using the original DEM. I then used the ArcGIS Hydrology toolset tools “Fill,” “Flow Direction,” and “Flow Accumulation” to recreate the channel network.

3.1.2.3 *National Elevation Dataset*

I also performed these basic routing and breaching processes on the USGS 1/3 Arc-Second Elevation Models. Like the LiDAR DEMs, these models include false obstacles that need to be breached. NED rasters come unprojected, mapped in a Geographic Coordinate System against the 1983 North American Datum (GCS North America 83). Spatial units are degrees. Before using this dataset for spatial analysis, I used the “Project Raster” tool in the Data Management Toolbox to redraw the DEM in the Universal Transverse Mercator Zone 18 (UTM18) coordinate system. This produced a raster with a resolution of 9.0678287 meters, giving each cell an area of 82.2255173 square meters. Using this value, I was able to convert the flow accumulation raster into upstream area values, and use that information in the regional regression curve calculations, described below. I did not perform any smoothing on the NED DEM.

3.1.3 **LiDAR Resampling**

While high-resolution digital elevation models can provide a great deal of useful information about a channel, they can also have resolutions that are so high they lead to their own problems. Specifically, LiDAR-derived elevation rasters may have cell sizes that are significantly smaller than the widths of the rivers being modeled. LiDAR is not able to pierce through the water surface well, and the conversion from the sparse LiDAR point-returns to a DEM results in a channel surface that is nearly flat, perhaps with some minor topography created by data “noise” rather than actual artifacts. Using the Hydrology Toolset in ArcGIS on this kind of DEM can create stream centerlines that meander and

drift randomly within the actual channel. The modeled channel, then, will be longer than the actual channel, and have a lower slope. Widths may also be challenging to extract from an overly-sinuuous channel, as cross-sectional lines drawn perpendicular to the modeled flow-path will often not be perpendicular to the actual centerline.

To address this issue, I used a method that was used and tested by Biron et al (2013). The method smoothes and reduces the resolution of the original, high resolution dataset with the goal of creating a DEM in which the channel is in the same location as in the original, but without the excess drifting and weaving of the modeled channel.

3.1.3.1 Resampling LiDAR in GIS

I first used the Focal Statistics tool to filter out errors, and then the Aggregate tool to increase the raster cell-size. Biron et al (2013) use a window size equal to about 10% of the width of the river in question when they run the Focal Statistics tool, and aggregate the data into cells sized at about 50% of the width of the river in question. I used the same

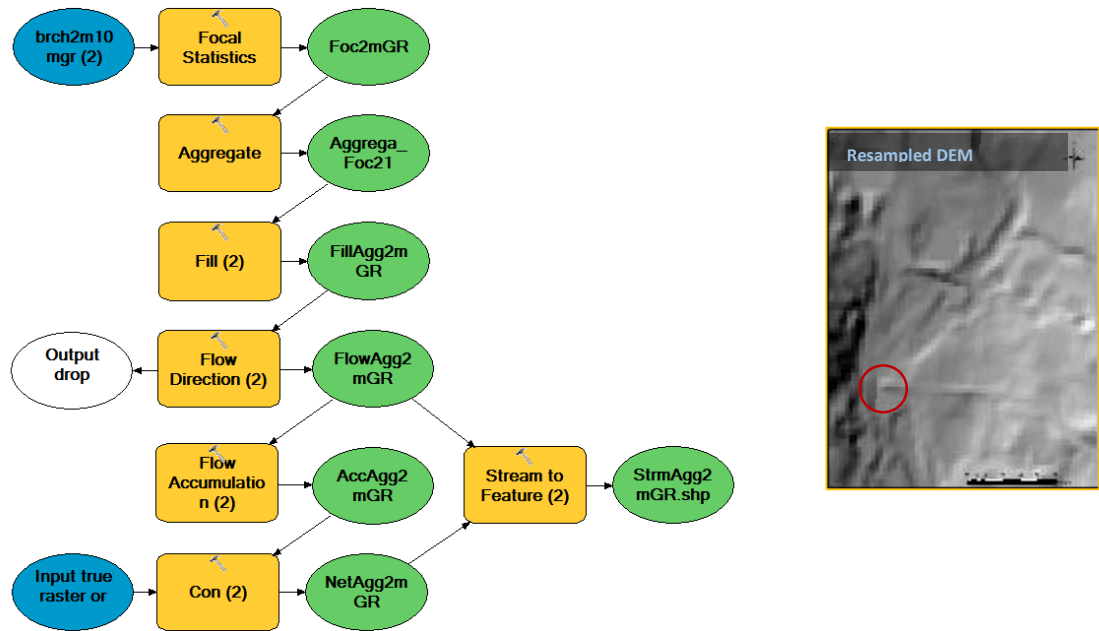


Figure 3-5: LiDAR Resampling, Routing, and Accumulation Work Flow Resampled DEM on Right

parameters, changing the values for each river I investigated. The rivers I explored were sometimes small enough that creating a window that was 10% of the river width for Focal Statistics tool would have created a window smaller than the cell-size of the raster. In those cases, I didn't run Focal Statistics.

I ran these tool on the DEM with obstacles already breached, and then used the flow routing and accumulation tools to generate a stream network based on the new, lower-resolution DEM.

3.1.4 Regional Regression Equations

A commonly-used method of estimating channel geometries and bankfull discharge values along a stream system is through the use of regression equations that relate those parameters to other physical attributes of a river or watershed. Typically these equations relate bankfull width, average bankfull depth, bankfull area, and bankfull discharge to the areal extent of the watershed above the channel cross-section in question. The values of these parameters, of course, are dependent on many more variables than simply upstream area - such as climate, geology, and land-use - so equations are usually developed for specific regions. The graphical representations of these equations are known as regional curves.

Regional curves have existed for the eastern United States for years (Dunne & Leopold, 1978), but the sites used for empirical observations of the relevant relationships are unknown, and so the accuracy of these equations in Massachusetts are questionable (Bent & Waite, 2013). In order to improve estimates, Bent and Waite (2013) recently developed regional curves for Massachusetts based on data collected from 33 sites around the state. These equations go beyond those previously available by accounting for variables other than upstream area. Analyzing the impact of different physical parameters

on channel geometries and discharges, those authors found that average basin slope is a significant factor along with upstream area. The curves they developed, then, relate channel geometries and discharges to both upstream area and average basin slope. These equations apply to streams where less than a quarter of the basin is urban, there is little flow regulation, the watershed area is between 0.6 and 329 square miles, and its mean slope is between 2.2 and 23.9 percent rise. All of the watersheds I investigated fall within these qualifications, although some of them have a certain amount of flow regulation.

3.1.4.1 *Regional Curve Calculations in GIS*

The equation for bankfull width developed by Bent and Waite (2013) is as follows:

$$\text{Equation 11} \quad W_{bkfl} = 10.6640 \times (A_{upstream})^{0.3935} \times (S_{basin\ mean})^{0.1751}$$

Here, W_{bkfl} is bankfull width, $A_{upstream}$ is the upstream drainage area in square miles, and $S_{basin\ mean}$ is the average percent-slope of the study basin. Upstream area, or the amount of land draining into a particular point, can be derived from a Flow Accumulation raster in ArcGIS. The value of a cell in a Flow Accumulation raster is equal to the number of cells whose flow drains into that particular cell. By converting that value into a measure of area by multiplying it by the spatial size of each cell, I was able to create a map showing the “Upstream Area” of each cell.

In order to determine the mean slope across the basin, I first had to isolate the basin of interest from the larger DEM. To accomplish this I used the “Watershed” tool in the Hydrology toolset, which creates a polygon of a watershed above a single “Pour Point” point shapefile. I used this polygon as a mask to extract the watershed of interest, used the “Slope” tool in the Surfaces Toolset to create a raster of the slope of each cell within

the watershed, and the “Get Raster Data” tool to find the mean slope value.

Once I had a raster of the upstream area and a value for the mean basin slope, I used “Raster Calculator” to plug this information into the regional curve regression equation. Initially, I calculated the base-10 logarithm of the bankfull width, and after derived the actual bankfull width, due to limitations of the Raster Calculator tool. Finally, I multiplied the results of these calculations by a raster of the channel centerline, where the channel had a value of 1 and everything else had a value of zero, creating a raster of the centerline where each cell had a value equal to the calculated width of the stream at that point.

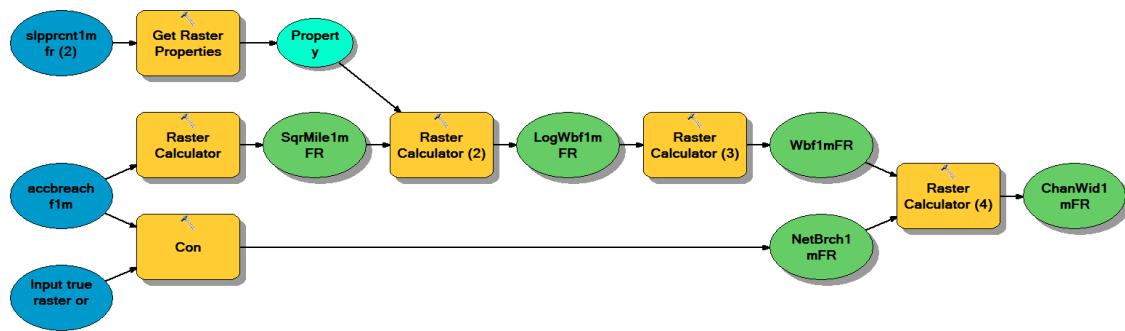


Figure 3-6: Regional Curve ArcGIS Workflow

3.1.5 HydroGeoMorphological Geoprocessing Toolset

The HydroGeoMorphological Geoprocessing (HGM) Toolset was developed by Guenolé Choné (Biron, Choné, Buffin-Bélanger, Demers, & Olsen, 2013), building on past work (Johansen, Tiede, Blaschke, Arroyo, & Phinn, 2011; Vocal Ferencevic & Ashmore, 2012). This toolset includes the tools “Binary River,” “River Width,” “Extract Data,” “Channel Slope,” “Flow Length from Point,” and “Breach.”

The “Binary River” tool demarcates the area of a river’s bankfull channel as a raster, where 0-values are inside, and 1-values are outside, the bankfull channel. The script begins with the channel centerline, created using the flow routing and accumulation methods

discussed. At each pixel marked as the centerline, the script analyzes pixels that neighbor the centerline pixel and are perpendicular to flow. If those pixels are below user-defined threshold values of changing slope and elevation, they are marked as within the bankfull channel. The process then continues with the next set of

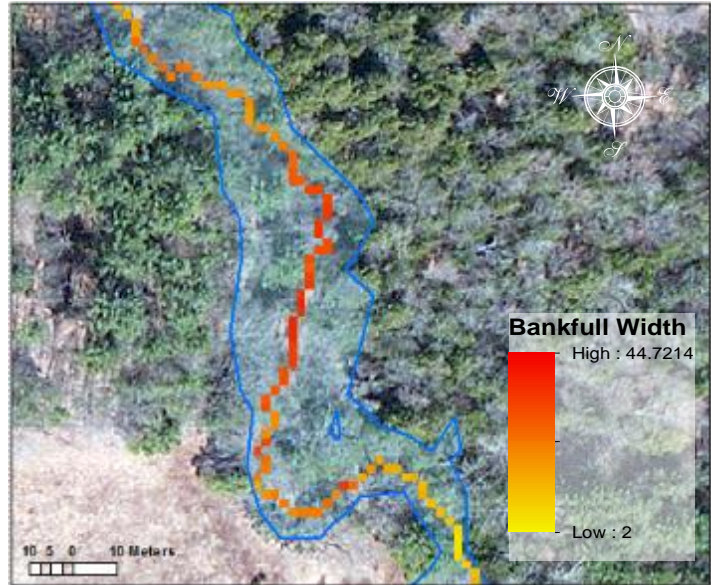


Figure 3-7: Binary River Output
Centerline pixels colored based on width at that point (from "River Width" tool). Delineated channel extent shown as blue polygon.

neighboring pixels, until the script finds pixels that are above the threshold values. This tool requires pre-existing elevation, slope, and flow direction rasters, as well as point shapefile where the script begins its analysis. The other inputs are values for elevation and slope threshold, and an iteration significance threshold. This last input defines a fraction value such that when less than that fraction of pixels being modeled by the tool are modified between two iterations, the tool stops.

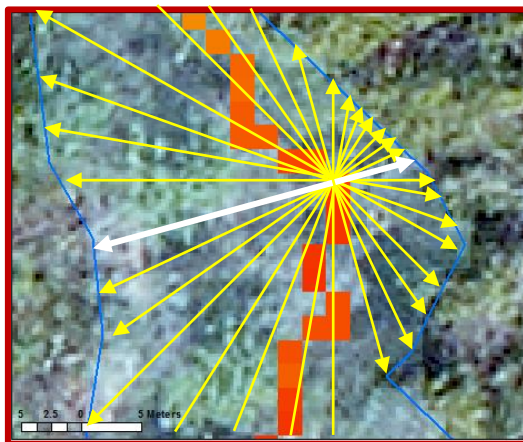


Figure 3-8: River Width Tool Schematic
Sixteen directions of bank-to-bank width are shown, shortest is highlighted in white.

The "River Width" tool turns the binary river raster into cross-sectional width information. At each pixel along the calculated centerline, the script measures the width of the river in sixteen directions. It then uses the smallest of those values as the bankfull width of the river, perpendicular to flow, at that pixel.

The "Extract Data," "Channel Slope," "Flow

29

Length from Point,” and “Breach” tools were not used in this research. “Extract Data,” based on scripts created by Vocal-Ferencevic and Ashmore (2012), is used to create a longitudinal profile of the river being investigated. “Channel Slope” is used to calculate the slope, in the direction of flow, of each pixel along the river. “Flow Length from Point” measures the distance, following the mapped channel, from the river head. The “Breach Tool” is similar to the “ToBreachThroughBridges” tool described earlier, though one must already have a map of the channel centerline to use it.

3.1.5.1 *HydroGeoMorphological Geoprocessing Toolset in GIS*

To map river width using the HGM Toolset, I ran the “Binary River” tool and then the “River to Width” tool, which uses the Binary River results as an input. “Binary River” uses rasters of elevation, slope (in degrees), and flow-direction, as well as a point shapefile designating a starting-pixel for its calculations. As the elevation rasters I used the LiDAR DEMs that had not been breached or filled. Slope rasters were generated using that same unprocessed DEM and the Spatial Analyst toolbox’s “Slope” tool, and the flow-direction rasters were generated from the DEMs that were breached and filled. This input data was chosen to produce a channel that follows the flow paths of the breached and filled DEM while producing channel width values based on the unaltered LiDAR data. Excess sinuosity within the channel caused by using the raw LiDAR data is not an issue in this case, because I am not using this tool to calculate channel length or slope.

The “Binary River” Tool also requires threshold values for elevation and slope change to be assigned in order to define stream banks. Biron et al (2013) used 0.5 meters as the threshold value for elevation change, and 12 degrees as the value for slope change. I used these same values, and wasn’t satisfied with the initial results. Therefore I ran the Binary Tool multiple times for each site, changing these threshold values to see how they affected the results. Each run of the tool was assigned a number, and those presented in

this report are HGM1 through HGM5. Elevation-change threshold values ranged from 0.50 to 1.00 meters, and slope-change threshold values range from 12 to 20 degrees. Parameters for each run of the HGM tool are listed in Table 3-1.

For each run of the Binary River tool I used 0.001 as the iteration significance threshold, meaning that the algorithm stopped processing when less than 0.1% of pixels were modified between iterations. I then used the output raster of this tool with the same flow-direction raster and starting-pixel shapefile to run the River-to-Width tool.

Table 3-1: HGM Tool Parameters

HGM Run	Elevation Threshold	Slope Threshold	Iteration Threshold
HGM1	0.50 m	12 degrees	0.001
HGM2	1.00 m	12 degrees	0.001
HGM3	0.50 m	20 degrees	0.001
HGM4	1.00 m	20 degrees	0.001
HGM5	0.75 m	15 degrees	0.001

3.2 Manual Analysis

“True” centerline locations either were not included in the field data I used, or did not have a high level of accuracy. Therefore, in order to have a “true” centerline against which I could measure the DEM-derived data, I delineated centerlines by hand using both aerial photography and LiDAR DEMs. In order to maintain consistency I drew all centerlines using the ArcGIS editing toolbox while viewing the map at the 1:1000 scale,. The lines I drew lie in the center of what I visually determined to be the bankfull channel, rather than along an inferred thalweg or in the center of the active channel.

These hand-drawn centerlines were created through thorough inspection of photographs and DEMs, but still contain a degree of uncertainty. In order to determine the significance of this uncertainty, I delineated the centerlines twice more for each watershed. Differences between these hand-drawn lines and the original set were averaged. This information

provides context for analysis of the significance of errors produced through the models, and are presented in the results section.

3.3 Overlaying Field Data

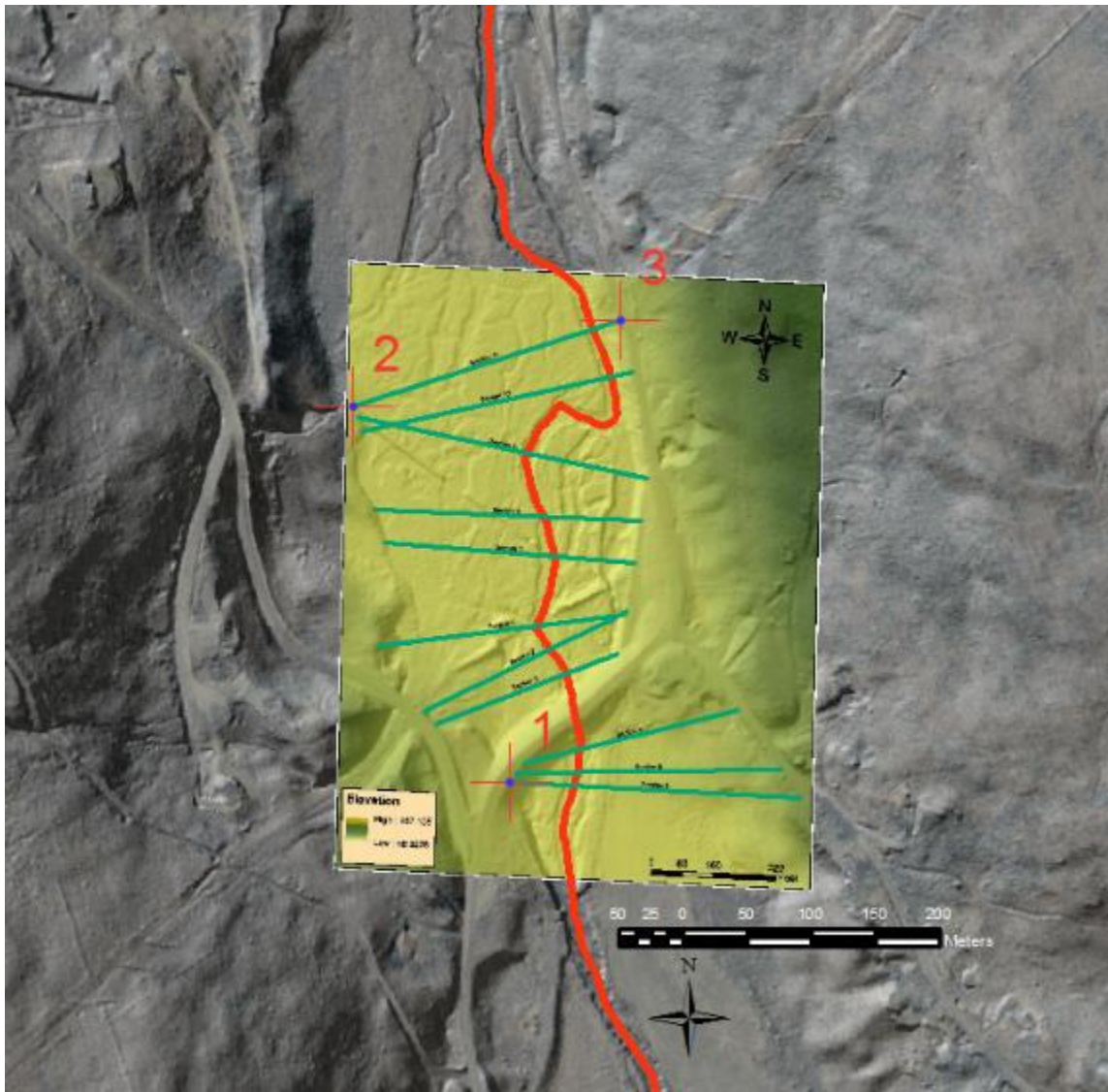


Figure 3-9: Overlaying Field Data

While some of the field-data provided for this project was georeferenced, other data was provided only as images of maps. In order to perform a meaningful comparison of this data to that produced from the DEMs, I used the Georeferencing toolset in ArcGIS to overlay and orthorectify the map images on top of the GIS map. I then used the Editor

Toolbar to draw new cross-section lines over the image, and the Interpolate Shape tool in the 3D Analyst toolbox to create a cross-sectional profile.

Additionally, I manually measured channel widths at specific locations using aerial photography and LiDAR separately, using the ArcGIS measuring tool to do so. When analyzing aerial photographs, I looked for changes in color between the immediate bank and farther sections of terrain, lighter-colored sediment bars, and the presence of trees. Where tree cover or shadows prevented reasonable visual assessment, “No Data” was marked. When assessing LiDAR data, I used the “hillshade” effect to make it easier to visualize changes in bank shape. I looked for changes in slope as represented by changes in shading, either from dark to light or light to dark depending on the aspect of the slope. I used those changes in shading to find the top of the channel banks, or the point at which flood waters would spill out of the main channel, to mark the edges of the bankfull channel.

In order to improve the robustness of this visual assessment method, for each site I took three measurements: one just upstream of the field-collected cross-section, one just downstream, and on along the same line. I averaged these three values and used that average as the bankfull value.



Figure 3-10: Manual Width Measurement
Longer purple lines show field-measured cross-sections. Shorter white lines show manually estimated widths.

3.4 Analysis Products

The key parameters of interest to me were channel planform shape and bankfull width. In order to assess the effectiveness of the methods I used, I created three metrics, as follows.

3.4.1 Sinuosity Index

Sinuosity index is the length of the channel divided by the length of the valley,

Equation 12
$$S = \frac{L_{\text{channel}}}{L_{\text{valley}}}$$

A straight stream has a sinuosity value of 1, while a highly sinuous stream will have a higher sinuosity value. Valleys themselves generally exhibit a certain amount of curvature and sinuosity, and so deciding where to demarcate valley ends is important for discerning smaller-wavelength sinuosity. The Deerfield River sub-basin field data included pre-mapped reach breaks, and so for those sites I calculated sinuosity as the length of the channel within each reach divided by the straight distance between each

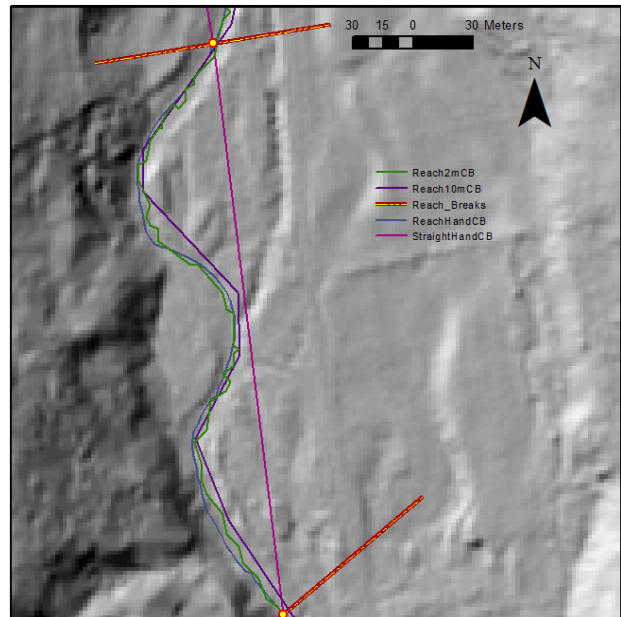


Figure 3-11: Sinuosity Index Calculation
Red lines are reach breaks. The straight purple line shows valley length. Different-colored meandering lines are modeled channel centerlines. Different centerlines have different shapes, and therefore different lengths and sinuosity index values.

reach break. Reach breaks were drawn in ArcGIS as straight lines perpendicular to the river. Because the different channel centerlines cross that reach polyline at slightly different locations, I measured the “straight” reach lengths for each version of the derived centerline from the point at which that particular centerline crossed the reach break. This

meant that the straight distance for the same reach had different values for each centerline.

No reach-break data was provided for the Fenton River. Based on the locations of the field-collected data, I selected a small stretch of river for study, and drew a polygon shapefile to demarcate the general valley setting. I used the "Polygon to Centerline" tool created by Tom Dilts of the Great Basin Landscape Ecology Lab (Dilts, 2011) to derive a valley centerline from that valley polygon. Finally, I used the "Transect Tool" (Ferreira, 2014) to draw reach breaks perpendicular to that centerline every one kilometer, and used those reach breaks to divide the stream into segments for sinuosity calculations.

3.4.2 Centerline Offset

The ArcGIS polylines representing these centerlines cross one-another in some places and lie next to one another in others. Combining multiple polylines creates a series of polygons with the polylines as their borders. The area of value of each polygon divided by the length of the measured centerline segment bordering that polygon gives the value of $\epsilon_{\text{centerline}}$, the average distance of that modeled centerline from the measured centerline:

$$\text{Equation 13} \quad \epsilon_{\text{centerline}} = \frac{A_{\text{polygon}}}{L_{\text{measured}}}$$

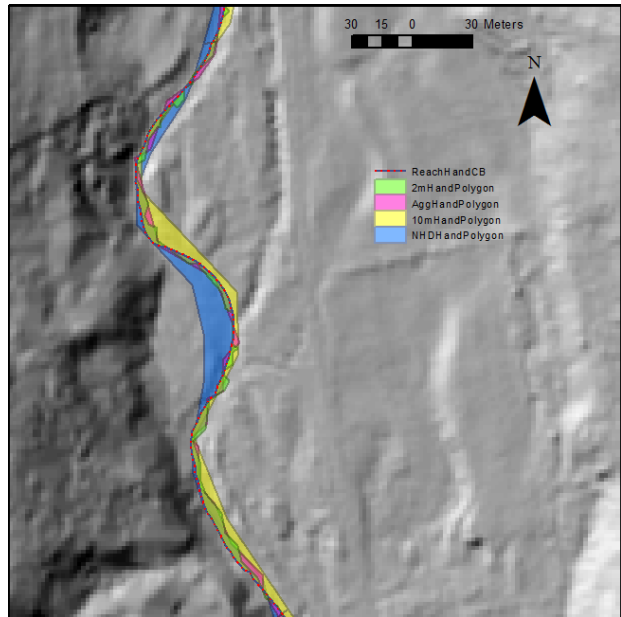


Figure 3-12: Centerline Offset Calculation
Colored polygons are those formed by combining modeled centerline polylines with hand-measured centerline polyline (blue and red dashed line)

I used the hand-drawn centerline based on aerial photos as the true centerline. I used the ArcGIS “Feature-To-Polygon” tool in the Data Management Toolbox to combine the hand-drawn centerline to the centerlines derived using the LiDAR, resampled-LiDAR, NED DEM, and the centerline created by the NHD. I drew and combined an additional bounding polygon so that polygons would be created at the up- and down-stream ends of the study area where the centerlines did not cross one another.

3.4.3 Width

For each watershed I had a select number of field-collected cross sections. Some of these included bankfull width values from those sites, and for others I determined bankfull widths based on the cross-section shapes. I then compared the model-derived bankfull width values for the channel at those cross-section sites against the field-based values.

CHAPTER 4

RESULTS

4.1 Sinuosity Index Error

4.1.1 Introduction

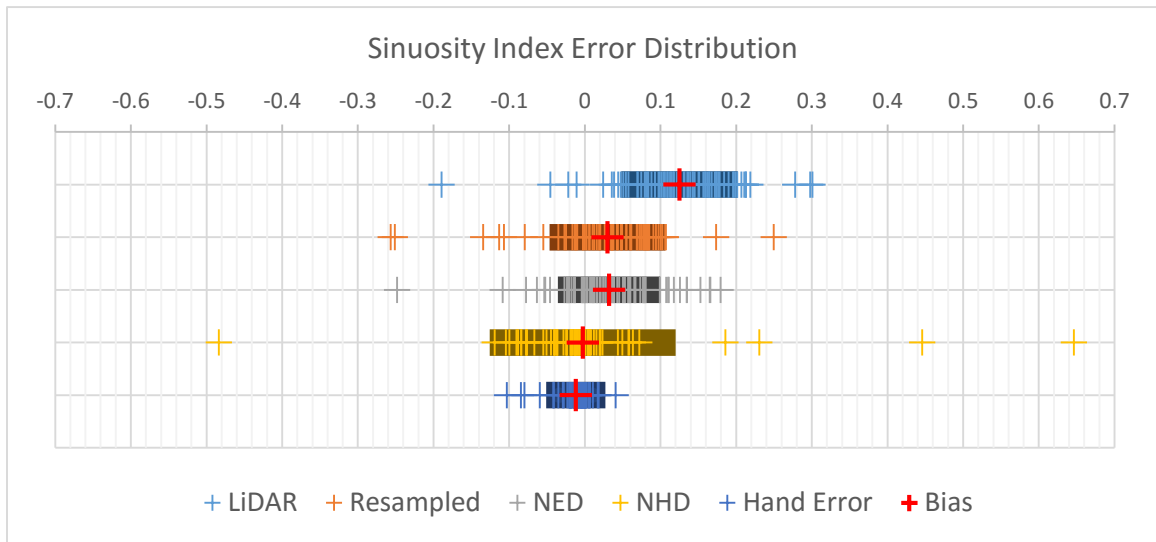


Figure 4-1: Sinuosity Index Error Distribution
Explanation in Text

Figure 4-1 shows the results of comparing the sinuosity values of modeled reaches to those of manually delineated reaches. Sinuosity index error is defined as the sinuosity of the modeled centerline at a particular reach minus the sinuosity of the measured centerline at that reach, which works out to:

Equation 14
$$\epsilon_s = \frac{L_{\text{modeled}} - L_{\text{measured}}}{L_{\text{valley}}}$$

An ϵ_s value of zero indicates a modeled centerline of the same length as the measured.

A positive ϵ_s means the modeled centerline is longer than the measured, most likely because of centerline drift and excessive meandering within the bankfull channel. A

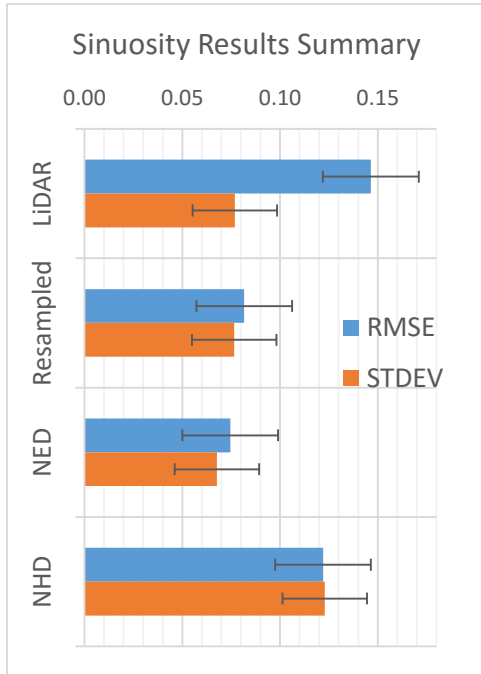


Figure 4-2: Sinuosity Index Error Summary
Explanation in Text

negative ϵ_s corresponds to a modeled centerline that is shorter than the measured due to the cutting off of measured meanders and other planform details. To give context to these values, measured sinuosity index values in this dataset range from just above 1.0 to nearly 3.0.

Each plus-sign in Figure 4-1 represents the ϵ_s value of an individual reach, with each row corresponding to a single centerline modeling method. The red plus-signs show the average ϵ_s derived from each modeling tool, or the bias in the

sinuosity results. Shaded areas show one standard deviation above and below this bias, such that the wider the shaded area, the greater the variance in the results.

Figure 4-2 again shows the standard-deviation of the sinuosity index error results, here in bar-graph format. Also included here is the Root-Mean-Squared Error (RMSE) of these data, where

Equation 15
$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^n (S_{\text{modeled},i} - S_{\text{measured},i})^2}{n}}$$

RMSE gives information about the overall accuracy of each model in a way that incorporates both the bias and the standard deviation of the errors.

Finally, Figure 4-3 shows modeled centerline sinuosity index results at each reach plotted against the sinuosity index values of the measured centerlines for the same reaches. In this graph, each point represents a single reach, and is color-coded based on the modeling

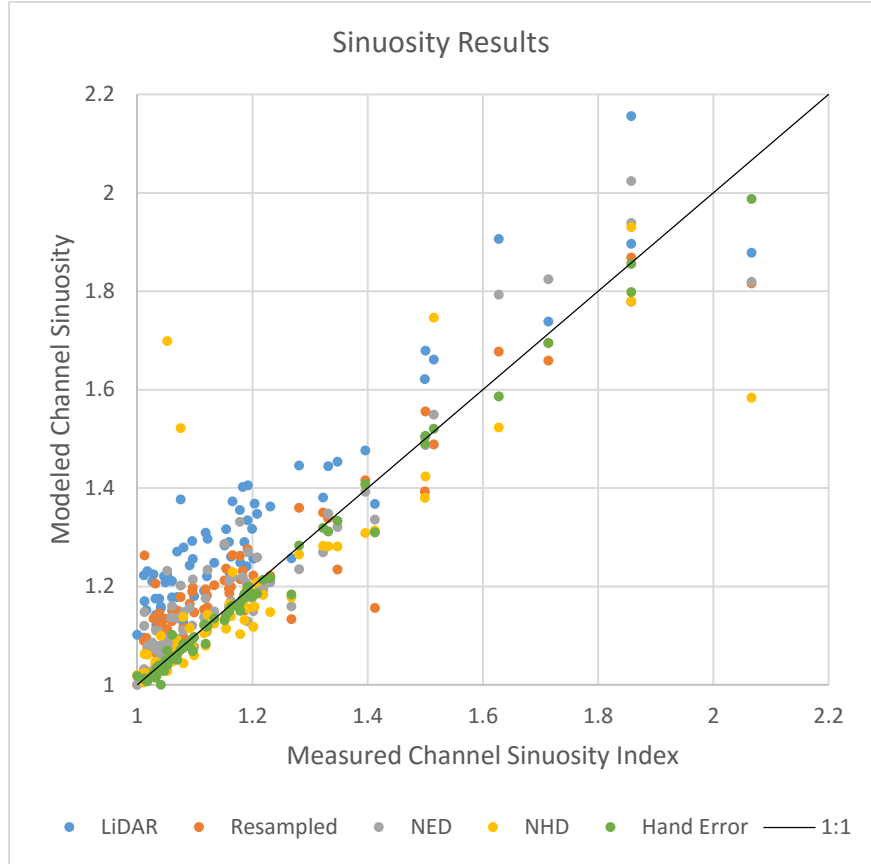


Figure 4-3: Chart of Modeled Sinuosity over Measured Sinuosity
Explanation in Text

method. The solid black line demarcates the one-to-one (1:1) line along which modeled sinuosity ratios that exactly-match measured ratios would lie. A point that falls above this line has a modeled sinuosity ratio higher than that determined visually for that reach, and one that falls below the line has a modeled sinuosity ratio lower than that measured visually. Each set of model results has a best-fit trendline (not shown in the graph) that is useful for analyzing relationships between sinuosity error magnitude and the magnitude of measured channel sinuosity index. The slopes of these trendlines are listed in Table 4-1.

Recall that centerlines were manually delineated twice more after the initial manual delineation, and those secondary lines compared to the original in order to provide information about the uncertainty of the baseline centerline. Figure 4-1 presents this

Model	Bias	St.Dev.	RMSE	Slope
LiDAR	0.12	0.08	0.15	0.92
Resampled	0.03	0.08	0.08	0.85
NED	0.03	0.07	0.07	0.97
NHD	0.00	0.12	0.12	0.90
Hand Error	-0.01	0.02	-0.02	0.97
Uncertainty	±0.01	±0.02	±0.02	±0.03

Table 4-1: Sinuosity Index Results Summary

uncertainty as a separate row titled “Hand Error” (bottom row). Figure 4-2 includes this information as error bars on the results for each model. Figure 4-3 includes the average sinuosity index values of the secondary hand-drawn centerlines as its own dataset. Table 4-1, which summarizes the results of key statistical measures, includes a row describing the difference between the secondary and primary hand-drawn centerlines. Uncertainty within each measure is taken to be plus-or-minus the absolute values of these results. The significance of differences between model results that fall within these uncertainty values must be questioned.

4.1.2 Results

The sinuosity ratios of modeled centerlines tend to be greater than those of measured centerlines, with the exception of the NHD lines. LiDAR results have the largest positive bias, with an average sinuosity error of 0.12. Put another way, centerlines generated by running the flow accumulation model on LiDAR DEMs are, on average, 12% more sinuous than measured channel centerlines. Centerlines created from both resampled-LiDAR data and from NED DEMs have an average error of 0.03. NHD results show no average error. See Figure 4-1 and Table 4-1.

This average error data provides information about the skew of model results, illuminating model biases. Standard deviation values provide information about the variance of modeling errors, while Root-Mean-Squared Error (RMSE) values capture both error magnitude and variance. Channel delineation using the LiDAR elevation models has the highest RMSE at 0.15, followed by NHD at 0.12, then resampled-LiDAR at 0.08 and NED at 0.07. This shows that LiDAR-derived- and NHD-centerlines have fairly large absolute error magnitudes while Resampled-LiDAR and NED-derived lines have relatively small absolute error magnitudes. NHD-produced centerlines have the highest variance in sinuosity errors, with a standard deviation of 0.12. LiDAR and resampled-LiDAR have standard deviations of 0.08, and NED error standard deviation is 0.07. This data reveals the poor performance of NHD results, evidenced by high RMSE and high standard-deviation, despite having an average error of zero. LiDAR-derived channels, on the other hand, have large sinuosity errors with a low standard deviation. Errors are negative on only three reaches, showing a degree of consistency in the mechanism by which those errors are created (excess sinuosity within the channel). Centerlines delineated from both resampled-LiDAR and NED data have low average errors, RMSE values, and standard deviations. See Figure 4-1, Figure 4-2, and Table 4-1.

In every modeled channel, sinuosity-error values tend to decrease (increase in the negative direction) as measured sinuosity values increase (Figure 4-3, Table 4-1). This pattern indicates a shift in the dominant form that sinuosity errors take. On relatively straight reaches, the models tend to produce centerlines that meander within the channel banks, although the degree to which that occurs varies by model. On more sinuous reaches, the models tend to delineate centerlines that do not fully capture measured meanders, cutting-off some of the river's curves. Figure 4-4 and Figure 4-5 show these two different forms of error. Reach 2 of the Green River (Figure 4-4) has a measured sinuosity of 1.01, but the modeled centerlines each show different amounts of noise and misplacement. The LiDAR-derived sinuosity is 1.17, resampled-LiDAR-derived is 1.26, NED is 1.15, and NHD is 1.06. Reach 7 of that river (Figure 4-5) has a measured sinuosity of 2.07, but the sinuosity index values produced by the LiDAR, resampled-LiDAR, NED, and NHD, are lower, at 1.88, 1.82, 1.82, and 1.58, respectively. Despite excess within-channel sinuosity

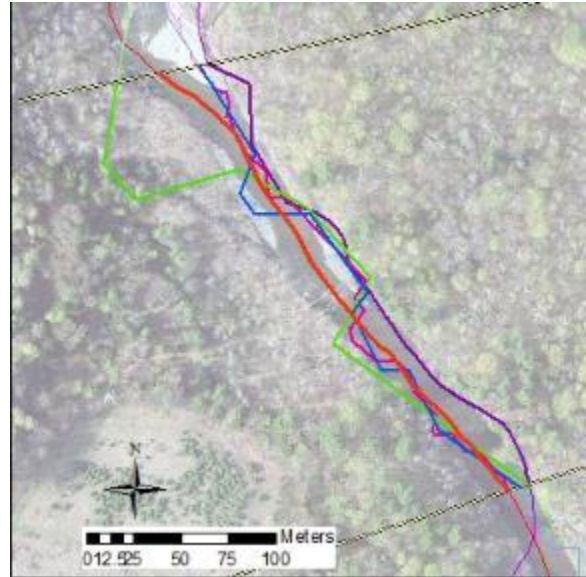


Figure 4-4: Excess Sinuosity in a Straight Reach Green River reach 2. Pink line is LiDAR-derived, green is from Resampled-LiDAR, blue is NED, and purple is NHD. Red is hand-drawn.

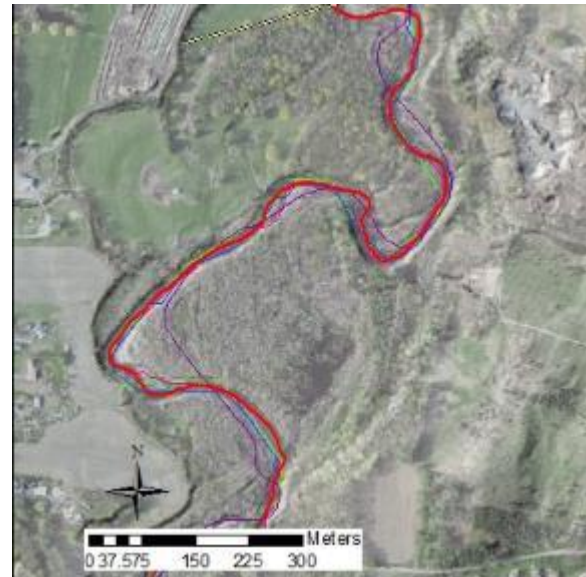


Figure 4-5: Deficient Sinuosity in a Sinuous Reach Green River reach 7. Pink derived from LiDAR, green from Resampled-LiDAR, blue from NED, purple from NHD. Red is hand-drawn.

that may exist, each channel centerline has been plotted in a path that cuts-off measured meanders to the point that all the modeled sinuosity index values are lower than those measured. Note that this trend, quantified by slope values lower than 1.0 in Table 4-1 is very minor in all models, with “Resampled LiDAR” being the only method that has a slope below 0.90. Slope value uncertainty is ± 0.03 .

Uncertainty in the mean sinuosity errors is ± 0.01 , and uncertainty in both RMSE and standard deviation of errors is ± 0.02 (Table 4-1). The difference in RMSE values between resampled-LiDAR and NED, as well as the standard deviation differences between all models outside of the NHD, are smaller than the uncertainty value of ± 0.02 . Those differences may therefore be insignificant.

Comparison between errors in the sinuosity ratios of modeled streams finds that resampling LiDAR data to increase pixel size decreases sinuosity error without having a significant impact on the variance of results. At the same time, NED-derived centerlines have a lower RMSE and a lower standard deviation than either LiDAR or resampled-LiDAR results. NHD lines, which have a very low error, also have a very high variance.

4.2 Centerline Offset Error

4.2.1 Introduction

Centerline spatial offset results give values in distance from the measured centerline. In order to perform meaningful comparisons between reaches of varying bankfull widths, I first normalized offset values to each reach’s “bank-distance,” defined as the distance from the measured centerline to the bankfull channel



Figure 4-6: Normalizing Centerline Offset
Red line is measured center, yellow is modeled.
White arrow is bankfull-width, orange arrow is "bank-distance," red line is centerline offset.

edge (Figure 4-6). This gives a value for “normalized offset error,” or “centerline-percent-bank-distance-error”

$$\text{Equation 16} \quad \delta_{\text{centerline}} = 100 \times \frac{\varepsilon_{\text{centerline}}}{\varepsilon_{\text{bank}}} = 100 \times \frac{A_{\text{measured-modeled polygon}}}{A_{\text{measured-bank polygon}}}$$

where $A_{\text{measured-modeled polygon}}$ is the area of the polygon formed by merging the modeled centerline to the measured centerline, and $A_{\text{measured-bank polygon}}$ is the area of the polygon formed by merging the bank-edge polyline to the measured centerline. These normalized spatial offset results are presented as percent bank-distance.

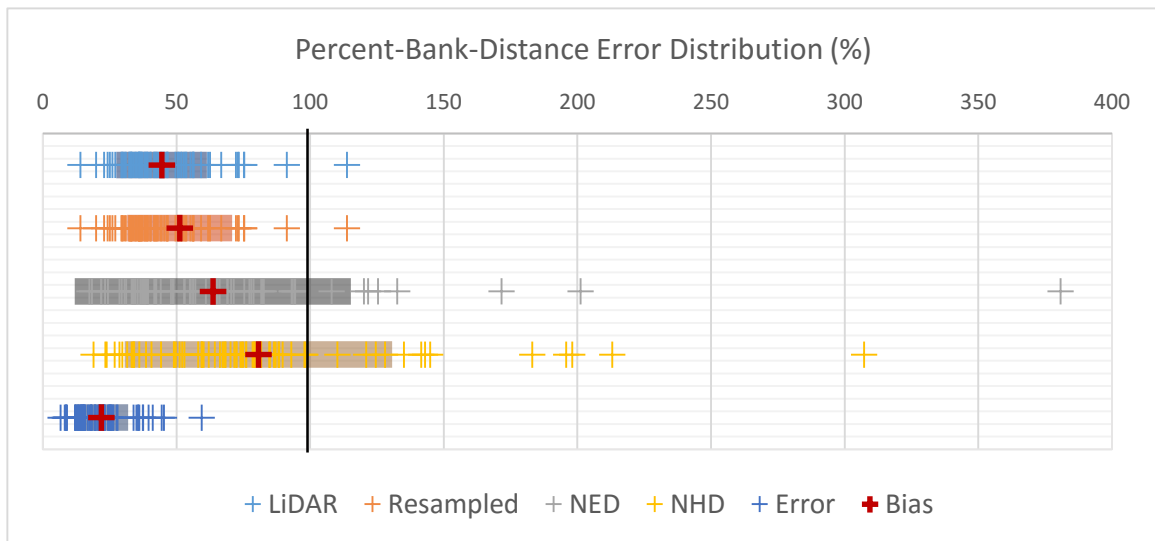


Figure 4-7: Percent-Bank-Distance Error Distribution

Bank edges were manually delineated for all of the streams being assessed. Using the same method described in section 3.4.2 (“Centerline Offset”) of creating polygons and dividing the area by the centerline length, I determined the average bank-distance for each reach. Modeled centerlines with an average error lower than 100 percent tend to be mapped within the measured channel, while those above 100 percent tend to fall outside the bankfull edge (Figure 4-7: black vertical line is at 100%). Uncertainty in mean error results, caused by variations in the manually delineated centerlines, is 21.9% of the bank-

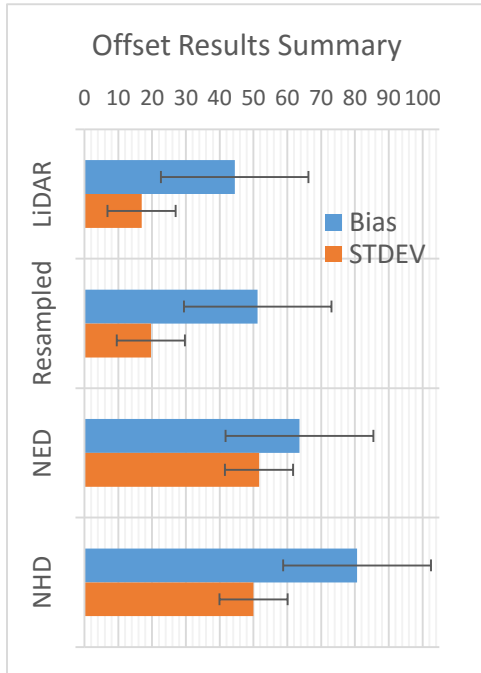


Figure 4-8: Offset Error Summary
Explanation in Text

distance. Standard-deviation uncertainty is 10.1% (See Table 4-2). This uncertainty is represented in Figure 4-7 as a separate row of results, and in Figure 4-8 as error bars.

4.2.2 Results

LiDAR-mapped centerlines have errors ranging from 14.0% to 113.8% of the bank-distance, with an average error of 44.4% and a standard-deviation of 16.8%. Note that the lowest error is less than the uncertainty inherent to the manual delineation of centerlines. Centerlines derived

from resampled-LiDAR DEMs have a lower error-bound of 24.3% and an upper bound of 106.9% of the bank-distance. The mean resampled-LiDAR offset error is 51.2% and the standard-deviation is 19.6%.

While centerlines derived directly from LiDAR have a lower average error and variance than those derived from resampled-LiDAR, it is important to recall that the differences in these results are lower than the uncertainty of the manually-delineated centerlines, and therefore cannot be compared with confidence. However, these two sets of results are clearly different than those from NED- and NHD-derived centerlines.

Centerlines produced from NED DEMs have a minimum error of 17.3% and a maximum error of 380.8% of the bank-distance. The average

Table 4-2: Offset Error Results Summary

Model	Bias (%)	St.Dev. (%)
LiDAR	44.4	16.9
Resampled	51.2	19.6
NED	63.6	51.7
NHD	80.6	50.0
Hand Error	21.9	10.1
Uncertainty	±21.9	±10.1

error and standard-deviation of this dataset are 63.6% and 51.7% of the bank-distance, respectively. NHD centerline errors have a minimum of 18.9%, a maximum of 307.24%, an average of 80.6%, and a standard-deviation of 50.0%.

These results show a trend of higher-resolution datasets producing centerlines with lower spatial errors and lower variance in these errors.

The differences in the average percent-bankfull error results are all less than the uncertainty value of 21.9%, but the trend of increasing error remains clear. The differences in the standard-deviation results between LiDAR and resampled-LiDAR and between NED and NHD are negligible relative to the uncertainty inherent in the manually delineated centerlines, preventing meaningful comparison between the members of each pair of results. However, it is clear that LiDAR and resampled LiDAR results have significantly lower variance than the NED and NHD results.

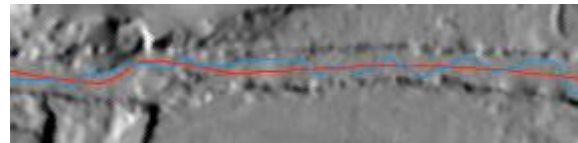


Figure 4-9: Erroneous Model within Banks
Red line is measured center, blue is modeled. Banks can be seen in hill-shade underlay. Fenton River.

It is worth noting that all of these results are below 100%, indicating that on average, all of these centerline delineation tools create lines that are within the measured channel extent. However, thirteen NHD-delineated reaches, eight NED-delineated reaches, three resampled-LiDAR-delineated reaches, and one LiDAR-delineated reach, have errors greater than 100% of the bank-distance.

Despite the baseline-uncertainty, and the fact that the average errors for all models are below 100%, it is clear that channels derived from NED and NHD are much less reliably placed within the channel extent. Centerlines derived from both LiDAR and resampled-LiDAR elevation models appear to successfully place the centerline accurately.

4.3 Bankfull Width Error

4.3.1 Introduction

The bankfull-width extraction models used in this research are able to map width values across an entire watershed. For the purpose of analysis, modeled bankfull-width results were compared to baseline data collected at specific sites in the field by researchers other than myself. Cross-sectional data for thirty sites distributed throughout the five study watersheds were provided for analysis.

Bankfull-width error ϵ_{width} is calculated as:

$$\text{Equation 17} \quad \epsilon_{\text{width}} = W_{\text{modeled}} - W_{\text{field-measured}}$$

Width errors were normalized to measured bankfull width values, such that error results are presented as:

$$\text{Equation 18} \quad \delta_{\text{width}} = \frac{\epsilon_{\text{width}}}{W_{\text{field-measured}}} = \frac{W_{\text{modeled}} - W_{\text{field-measured}}}{W_{\text{field-measured}}}$$

Statistical measures used to analyze width results include mean normalized error, standard deviation of the normalized error, and Root-Mean-Square of the normalized error, or:

$$\text{Equation 19} \quad \text{RMSNE} = \sqrt{\frac{\sum_{i=0}^n \left(\frac{W_{\text{modeled},i} - W_{\text{field-measured},i}}{W_{\text{field-measured},i}} \right)^2}{n}}$$

For the purposes of this analysis, field-derived bankfull width data is assumed to be accurate, and so no uncertainty values are associated with the data. This is reflected by the lack of an extra row in Figure 4-10 or in Table 4-3, or error bars in Figure 4-12.

4.3.2 Results

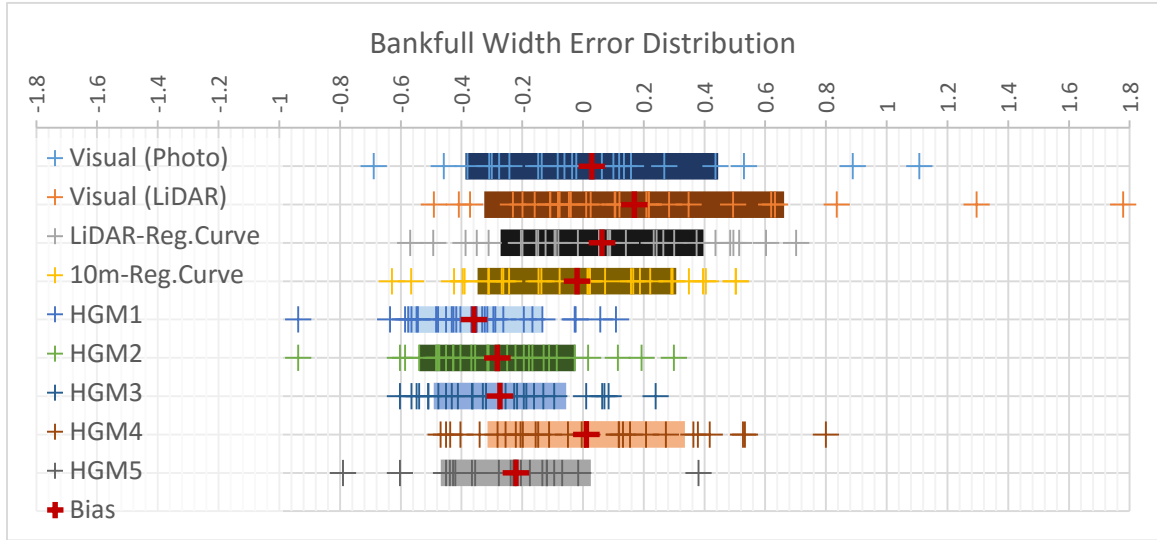


Figure 4-10: Bankfull Width Error Distribution

Recall that I used the Bent & Waite (2013) regression equation to calculate bankfull width values across the study watersheds using both the LiDAR and the NED elevation models. Results from the LiDAR tended to slightly overestimate bankfull width, with an average width estimate 6% greater than the field-measured width. NED-based regression equations produce width estimates that are on average only 2% less than the field-measured widths. Despite these low averages, the variance in the error is quite high, with standard-deviation values of 33 percentage-points for both models. RMSNE results are 0.33 and 0.32 (33% and 32% of bankfull width) for the LiDAR-based regression and NED-based regression, respectively. See Table 4-3, Figure 4-10, and Figure 4-12.

Bankfull-width results were also derived from the HGM tool, which was applied to the LiDAR DEM. Recall that the tool was run five times, each time with different input threshold values (Table 3-1). Runs were labeled as HGM1 through HGM5. Mean normalized error ($mean \delta_{width}$) values for these runs, beginning with HGM1 and moving to HGM5, are negative-0.36, negative-0.28, negative-0.27, 0.01, and negative-0.22. Standard-deviations are 0.22, 0.26, 0.22, 0.33, and 0.25. RMSNE results are 0.42, 0.38, 0.35, 0.32,

and 0.33. The HGM toolset tends to underestimate widths, with the exception of HGM4, which has an average error of nearly zero. The standard-deviations of HGM-tool errors are smaller than those of the regression

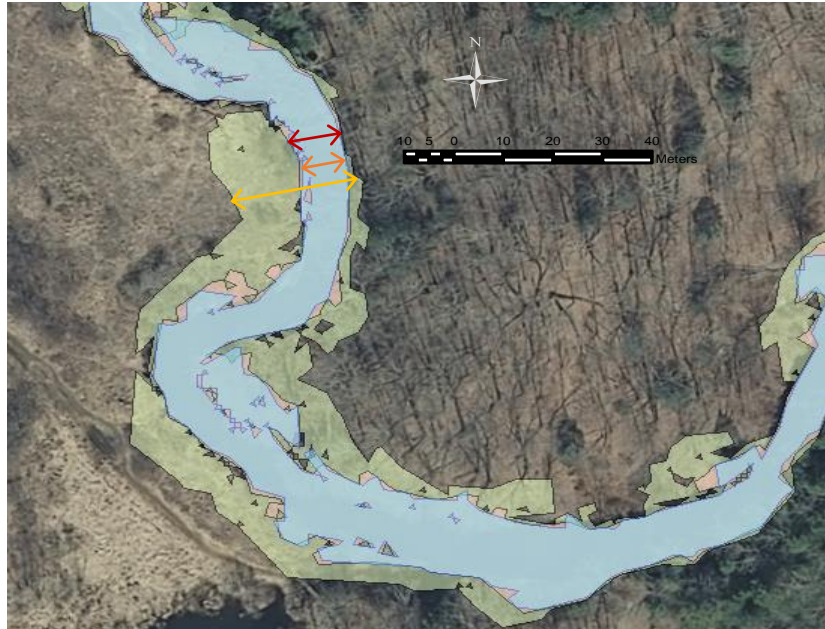


Figure 4-11: Example of Varying HGM Tool Results, Fenton Brook

equations, again with

the exception of HGM4, which has a variance of 0.33 (33% of measured bankfull). Important characteristics to note in the HGM tool results are the variation between one run of the tool and another, and the relatively low standard deviation values given by all of the HGM tool results when compared to regression equations and visual estimates.

Visual estimates of bankfull-width, using both aerial photography and LiDAR-DEMs with the hillshade-feature in use, show low accuracy. δ_{width} values are somewhat low (0.03 for photo-based and 0.17 for LiDAR-based estimates), but the standard-deviations of the errors (0.42 and 0.49), as well as the RMSNE values (0.41 and 0.51) are quite high, and indicate very poor reliability of this method.

There are a few key points to take away from these results. Bankfull widths estimated using the Massachusetts-specific USGS regression equations as well as visual inspection of aerial photographs and LiDAR DEMs have large errors in both the positive

Table 4-3: Bankfull Width Results

Model	Bias	St.Dev.	RMSNE	Slope
Visual (Photo)	0.03	0.42	0.41	0.78
Visual (LiDAR)	0.17	0.49	0.51	1.10
LiDAR-Reg.Curve	0.06	0.33	0.33	0.25
10m-Reg.Curve	-0.02	0.33	0.32	0.80
HGM1	-0.36	0.22	0.42	0.83
HGM2	-0.28	0.26	0.38	1.15
HGM3	-0.27	0.22	0.35	0.66
HGM4	0.01	0.33	0.32	0.98
HGM5	-0.22	0.25	0.33	0.86

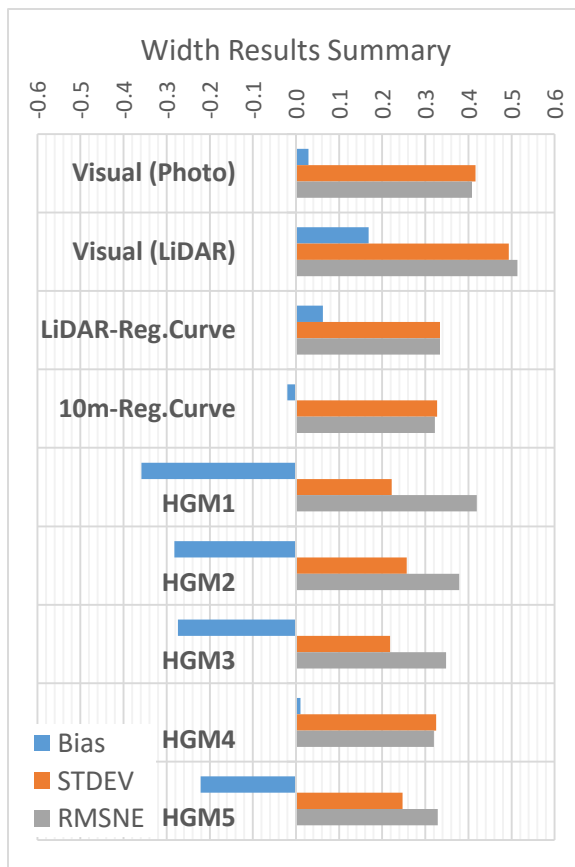


Figure 4-12: Bankfull Width Results Summary

(overestimate width) and negative (underestimate width) directions. Over- and under-estimates are about equal for each of these methods, such that the average error is very near zero. Using the HGM tool with a slope threshold of 20 degrees and an elevation threshold of 1.0 meter (HGM4) produces a similar pattern of erroneous estimates. The errors and variance results from the HGM4 run are on par with those from the regional regression equation analysis and are lower than those from visual estimates. The rest of the HGM tool runs have lower variability in percent-

error values, but also have a strong tendency to underestimate widths. It is clear, though,

that input thresholds to this tool have a significant impact on results, and it is possible that further adjustments would produce significantly different error distributions.

CHAPTER 5

DISCUSSION

The previous section presented feature-extraction results broken down into the three categories of sinuosity, centerline spatial offset, and bankfull width. Recall that the first two categories provide information on the accuracy of model-extracted centerlines, with sinuosity serving as a proxy for centerline planform shape and centerline spatial offset serving to quantify centerline placement. The third category covers the accuracy of model-extracted bankfull width values.

5.1 Centerline Mapping

5.1.1 Centerline Discussion

Reframing sinuosity and centerline spatial offset as two indicators of a single capability (centerline extraction) allows for analysis of each tool's competence with regard to that capability. Thus it can be observed that flow routing and accumulation using a LiDAR DEM produces a centerline that captures the measured centerline location while poorly modeling its shape, specifically due to excessive meandering and drift within the channel extent. Running flow accumulation after first resampling the LiDAR DEM using the Biron et al (2013) methodology creates a centerline that retains the spatial-location accuracy of the original LiDAR-derived line while also improving planform shape accuracy through reduction of excess meanders. Applying the flow accumulation method to the 1/3 arc-second NED elevation model produces a centerline with a planform shape that is as or more accurate than that extracted from resampled LiDAR, but with a slightly less accurate spatial location. Finally, NHD lines seem to match measured channel shape more closely than any of the DEM-based models, but with significant spatial offset errors.

5.1.2 Centerline Modeling Sources of Error

The errors found in the centerline modeling process can be the signals of a number of different problems with the data being assessed.

5.1.2.1 *Model Failure*

First, the modeling method exhibiting an error may function poorly at the reach in question. This can occur if there are multiple active channels or abandoned channels that “capture” the modeled flow, if the reach with an error is a marsh environment without a clearly formed channel in the elevation dataset, or if obstacles captured by the elevation dataset block the correct flowpath.

Reach 3 of Pelham Brook is one example that falls into this category (Figure 5-1). It consists of an extremely low-gradient wetland created by infilling of a dammed pond. No channel is visible in the LiDAR, and each modeled channel passes through a unique route. The curve as mapped from aerial photography is followed by the NHD channel, nearly followed by the NED and LiDAR channels, and completely cut off by the resampled-LiDAR channel.

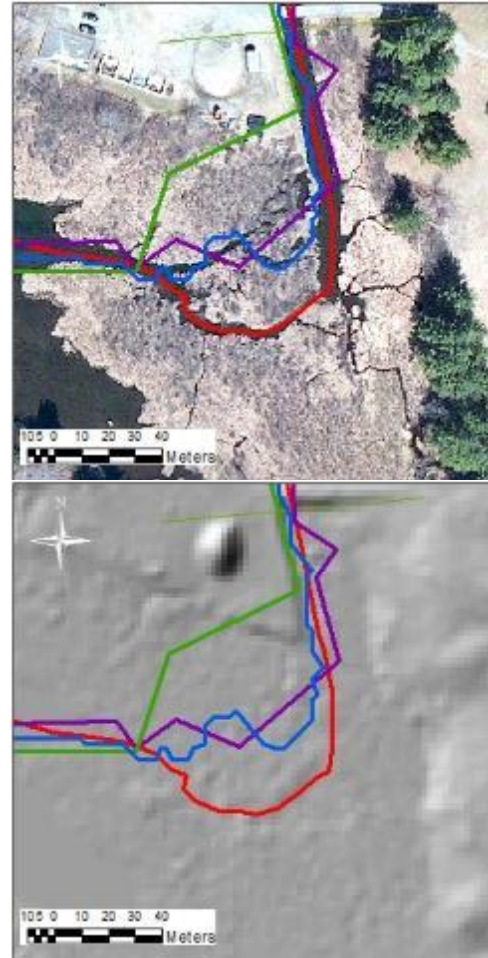


Figure 5-1: Model Failure
Hand-drawn centerline is red. LiDAR- (blue) and NED- (purple) lines follow secondary channel. Resampled-LiDAR (green) cuts curve off completely. NHD line not shown.

5.1.2.2 *Dataset Inconsistency*

Second, the model with an error may be based on a dataset from a different time period than the aerial photographs from which the measured centerline was derived. The channel may have actually moved between these two time periods, leading to a perceived modeling error. It's important to note that rivers evolve constantly, and so accurate mapping requires up-to-date information. A centerline modeling method that is extremely accurate but cannot be updated regularly due to data limitations, such as the expense of producing a LiDAR DEM, has a legitimate limitation. Therefore such errors should not be disregarded.

One such example occurs in the modeling of reach five of the Green River, which forms a low-gradient, forested, multi-thread stream. The sinuosity index error values of the modeled channels deviate significantly from the general trend at these reaches for all models. The deviation is clear to see at the downstream end of reach 5 and upstream end of reach 6. Examination of the aerial photographs and the digital-elevation-models at these sites reveal that the channel most

likely migrated drastically soon before the photographs were taken. The hand-drawn centerline, based on the aerial photography, was therefore depicting a very different channel than the models were capturing.

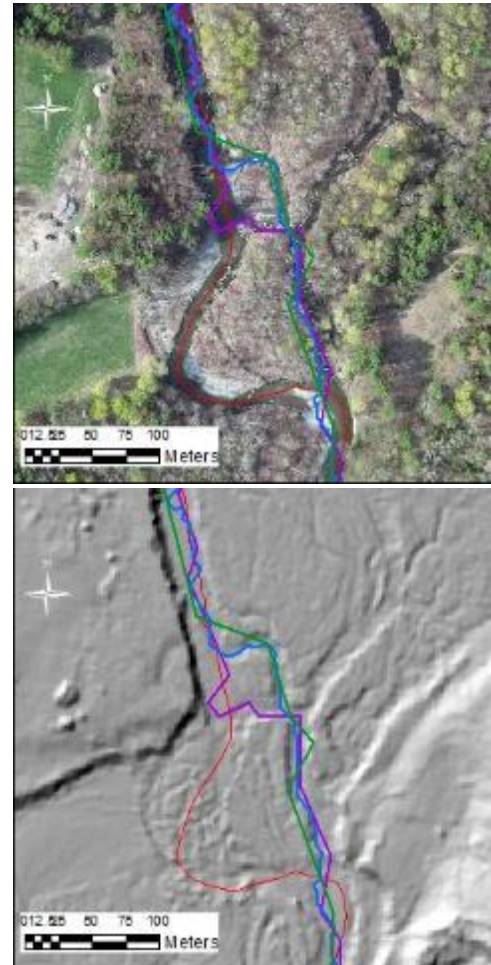


Figure 5-2: Dataset Inconsistency
Red, hand-drawn line follows the main channel as seen in the aerial photo (topt). LiDAR (blue), Resampled-LiDAR (green, and NED (purple) channels follow each-other fairly closely.

5.1.2.3 Operator Error

Finally, the error may be the result of mistakes made in the process of measuring the centerline or the bank edge. These errors do not represent weaknesses of the models in question, rather they indicate weaknesses in the model analysis method. The uncertainty inherent in the results because of human error is explained in the results section.

5.2 Bankfull Width

5.2.1 Bankfull Width Discussion

Visual estimates, which I expected to be closest to field-measured results when compared to the automated modeling tools, seem to have been the least-accurate method of determining bankfull width.

Average error was near zero for widths estimated from aerial photography, but the variance of the errors was very large. Widths estimated visually from LiDAR had a low average error when compared to other tools, but this error was not insignificant, and the variance in these results was also

very large. Visual analysis of both aerial photography and LiDAR

DEMs captured the extents of the widest channels more accurately than any of the models (Figure 5-3).

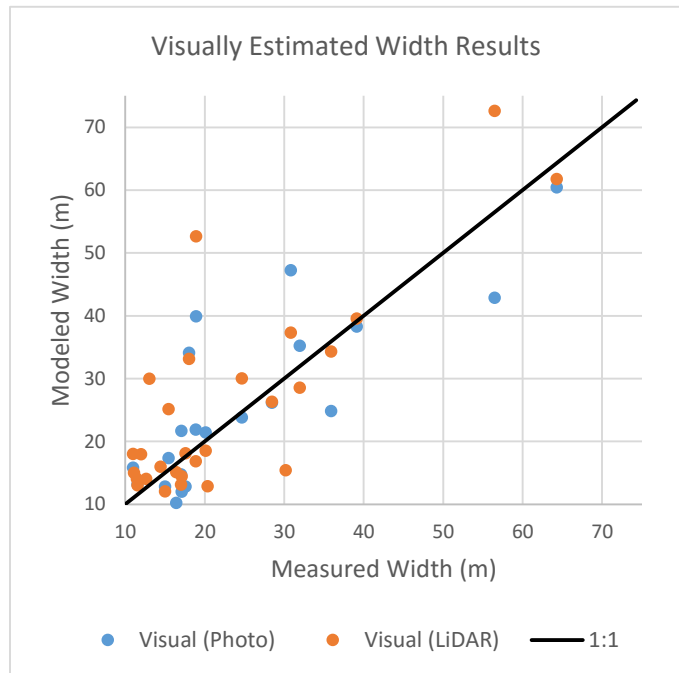


Figure 5-3: Visual Estimate Results Modeled Width over Measured Width

The regression equations that I used incorporate the average slope of the terrain of the watershed in question, giving them a degree of sensitivity to regional conditions. Beyond that consideration, this tool bases width estimates only on upstream area, and lacks any capacity to respond to small-scale, local features or processes. This deficiency is especially evident in the tool's underestimates of the

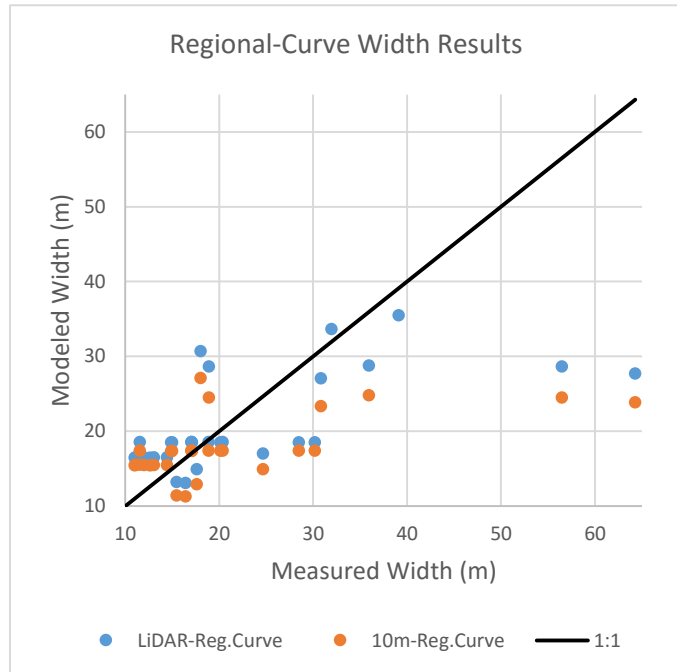


Figure 5-4: Regional-Curve Results: Modeled Width over Measured Width

widths of wider streams (Figure 5-4). The average error, on the other hand, is very near zero. It may be the case that these regression equations are able to model New England regional trends well (leading to the low average error), but that a more flexible tool is required to capture those localized variations (as seen in the high standard deviation in error). Finally, it is important to consider the fact that regression equations, which are derived empirically based on data chosen by particular researchers, vary, and other equations have been developed that may be more or less applicable to certain region. Other equations may have performed differently at these sites.

The HGM Toolset modeled local changes in width in ways that the Bent and Waite (2013) regression equation was not able, and over an extent that would not be feasible to examine visually. This tool tended to underestimate channel widths, especially on very wide streams (Figure 5-5). As mentioned in the Results section, there are two particularly interesting aspects of the results of this tool: the degree of variance in the errors generated

by the model, and the differences between the results of each run. Standard-deviations of the errors generated by every run of the HMG tool were lower than those created by application of the regression equation and by visual estimates. Despite highly negative mean error values, the lower variances indicate a degree of consistency in these tools that does not exist in the other methods used. The variability

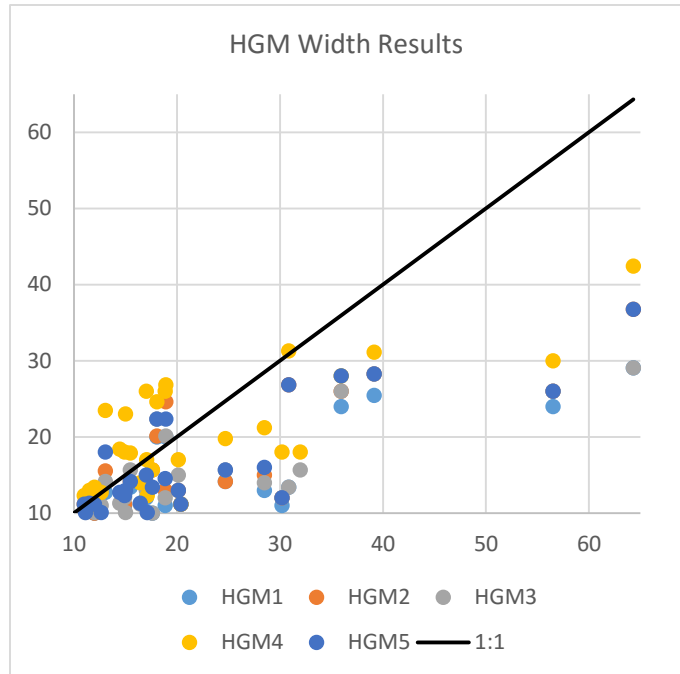


Figure 5-5: HGM Width Results: Modeled Width over Measured Width

between the results of the different runs, highlights the importance of the user-defined threshold values, and suggests the possibility of improvement to the tool's performance through alteration of those values. HGM4 results prove that this tool can perform at least as well as existing methods.

5.2.2 Bankfull Width Sources of Error

Extracting bankfull-width information from digital elevation models proved to be a challenge. In an alluvial system that exists in a state of relative equilibrium, bankfull stage is by definition associated with topographic features (specifically the top of the bank), but in complex real-world systems that may include erosive environments and disequilibrium caused by geologic, climatic, or anthropogenic forces, such topographic indicators may be missing or unreliable. The upland New England watersheds that were the focus of this research include erosion-dominated reaches as well as streams that continue to adjust in response to present human activities, past deforestation, and changes in climate and

geology since deglaciation. For this reason, determining bankfull flow information from topographic features, and consequently from digital elevation models, is a challenging task, to say the least.

A specific source of error arose in a low-gradient areas with multiple channels. Because the flow routing and accumulation tools I used created only a single centerline, bankfull width estimates based on both accumulation (regression equations) and bank-distance from the centerline (HGM tool, visual estimates), were unable to capture these more complex systems.

CHAPTER 6

CONCLUSIONS

6.1 Feature Extraction

6.1.1 Channel Planform

This research has found that basic flow-routing and accumulation algorithms in a GIS, which digitally simulate watershed-scale fluvial processes, can map a channel's planform shape and location using both high-resolution LiDAR and lower-resolution NED digital elevation models. With a minimal amount of preprocessing, specifically through the resampling of the DEM to create a coarser-scaled dataset, the LiDAR data can be used to model a channel that highly correlates to the shape and location of the mapped channel. NED-derived channels model the mapped channel shape with even greater accuracy, and model the channel location only minimally less accurately, than the resampled LiDAR. Given that these two elevation datasets produce very similar results, the decision about which to use in a given situation depends on factors other than model performance.

6.1.2 Bankfull Width

None of the tools studied in this thesis succeeded at extracting bankfull width values with a high degree of accuracy and consistency. Nevertheless, analysis of the LiDAR data by the HGM toolset and through manual visual inspection did capture details that could not be resolved using regression equations, the only automated analysis method used in this research that was applicable to the NED DEM. Variability in HGM results and the better performance of that tool on wide reaches than regression equations indicates potential for improvement to this tool with further research. With improvement, the HGM toolset or other

methods designed to extract bank information from LiDAR terrain models could prove to be powerful tools is fluvial geomorphology.

6.2 Further Work

6.2.1 Alternative Approaches

The automated extraction methods I tested are just a small selection of a fairly wide variety of tools that either are designed or are in the process of being designed to extract fluvial-geomorphic data from LiDAR. These tools likely have different strengths and weaknesses than those described and tested in this thesis, and exploring their capabilities would provide further insight into the general capacity of automated methods to evaluate geomorphic parameters from high resolution DEMs. The following are examples of two very different approaches to geomorphic feature extraction.

The Antonarakis et al (2008) method uses raw LiDAR data, rather than a DEM derived from LiDAR data or bare-earth LiDAR data, to differentiate between different kinds of landcover, including surface water. Using C++ programming language and ArcGIS and MatLAB, the authors use the pattern- of the elevations of different signal returns (first, last, and intermediate) and the percent of emitted signals returned (as opposed to absorbed) to determine the surface type at a given point. Using this method, they are able to delineate streams and other surficial water bodies, as well as different types of ground cover.

Giulia Sofia and Paolo Tarolli at the University of Padova in Italy, with other collaborators are working on a tool that calculates a parameter known as the “elevation percentile,” or E%, across a landscape, based on a DTM of that landscape. The elevation percentile value of a cell is calculated by counting the number of cells within a given neighborhood around the cell in questions that have a lower elevation than the central point, and dividing that number by the total number of cells within the window. This gives each cell a value

between zero and one, with zero meaning that the topography at that point is concave up, and a one meaning it is concave down. After this process, a QQ-plot is used to determine the E% value that indicates the bankfull height, allowing for bankfull-width channels to be mapped. Bankfull width can then be calculated by measuring width perpendicular to flow direction (person communication with Giulia Sofia and Paolo Tarolli, August 10, 2014).

6.2.2 Tool Refinement

Flow routing and accumulation, LiDAR resampling, regional regression analysis, and HGM toolset application can each be accomplished in different ways using different techniques, user inputs, and data sources. Automated extraction of geomorphic data from could potentially be improved by making different choices with regards to those techniques, inputs, and sources, while using the same basic tools and methods described in this thesis.

6.2.2.1 *Flow Routing and Accumulation*

The flow routing tool built into ArcGIS 10, and used in this thesis, uses the D8-algorithm described in the methods section. Tarboton (1997) created an alternative algorithm called D-Infinity. This algorithm models flow partitioning (water can flow in more than one direction out of a pixel) between up to two neighboring cells, effectively allowing for flow in an infinite number of possible directions. Other automated watershed delineation and channel designation models have been developed by Mark et al (1984), Band (1986), Moore et al (1988), Marz and Garbrecht (1993), and Miller et al (2002). The existence of this variety of flow routing, flow accumulation, and channel mapping methods shows that there are choices to be made even in this relatively simple aspect of DEM analysis.

6.2.2.2 *LiDAR Resampling*

In this thesis project, LiDAR DEMs were resampled to create a new raster with pixels approximately half the bankfull width of the river being studied. Of course, bankfull widths change as one moves downstream as well as throughout a watershed based on local conditions. The pixel value chosen for resampling is therefore somewhat arbitrary. Further exploration of resampling methods and resampled dataset pixel sizes could improve the quality of channels mapped in this way. It may be preferable to use a larger or a smaller pixel size, to choose the elevation values of redrawn pixels differently, or to change these parameters within a single basin. Being thorough in this process may improve results, but may also lead to increased time and effort for data processing

6.2.2.3 *HGM Toolset*

The HGM Toolset has great potential to improve the power of remotely-sensed elevation data in basin-wide river assessments. Its placement of the channel centerline uses the same calculations as the ArcGIS flow-routing algorithms, allowing for easy translation between the HGM-derived channel and watershed-data such as accumulation and slope.

Regardless of the threshold parameters used in this thesis research, the HGM toolset did not consistently produce width values matching those measured in the field. These estimates could be improved by further research into where the tool does work well under different elevation- and slope-threshold regimes, and then changing those threshold inputs depending on local factors. Guenole Chone (personal communication, May 21, 2014) was able to alter the tool input parameters such that one can now vary the slope and elevation threshold values based on an underlying raster. This allows a user to change those values based on local slope, surface geology, or land-use, for example.

Further exploration of the accuracy of the HGM toolset at predicting widths using different threshold inputs in different parts of a channel network could make this new capability very useful. Pursuing this exploration would be very beneficial.

6.3 General Conclusions

NED datasets are available free of charge across the entire United States, and are updated by the USGS as newer data is made available. LiDAR data covers a much more limited area, and collection of new data is expensive. On the other hand, as the LiDAR collection and processing industries grow, costs may drop and coverage increase. Furthermore, LiDAR data that does exist will typically be newer than any other available NED coverage for that area. LiDAR-derived elevation products have the data needed to perform watershed-scale process modeling that can resolve process and form at the resolution of specific sites, while coarser-scaled elevation data does not have that localized information. As improvements are made to data collection and processing methods, LiDAR technologies have the potential to provide an accessible source of key information for communities and decision-makers.

WORKS CITED

- Bagnold, R. A. (1966). *An Approach to the Sediment Transport Problem from General Physics*. United States Department of the Interior, Physiographic and Hydraulic Studies of Rivers. Washington: United States Government Printing Office.
- Bent, G. C., & Waite, A. M. (2013). *Equations for Estimating Bankfull Channel Geometry and Discharge for Streams in Massachusetts: U.S. Geological Survey Scientific Investigations Report 2013-5155*. Reston, VA: USGS.
- Biron, P. M., Choné, G., Buffin-Bélanger, T., Demers, S., & Olsen, T. (2013). Improvement of Streams Hydro-Geomorphological Assessment Using LiDAR DEMs. *Earth Surface Processes and Landforms*, 38, 1808-1821. doi:10.1002/esp.3425
- Brierley, G. J., Fryirs, K., & Cohen, T. (1996). *Geomorphology and River Ecology in Southeastern Australia: An Approach to Catchment Characterisation. Part one: A Geomorphic Approach to Catchment Characterisation*. Working Paper 9603, Macquarie University, Graduate School of the Environment, Sidney, Australia.
- Brink, J., Skewes, W., & Henry, E. (2013). *Geomorphologic Characteristics of the Fenton River: UConn Forest Meanders*. Prepared for Geography 5310, University of Connecticut.
- Carbonneau, P. E., & Piégay, H. (2012). The Growing Use of Imagery in Fundamental and Applied River Sciences. In P. E. Carbonneau, & H. Piégay, *Advancing River Restoration and Management, Volume 1: Fluvial Remote Sensing for Science and Management* (pp. 1-18). Chichester, West Sussex, UK: John Wiley & Sons, Ltd.
- Cavalli, M., Tarolli, P., Marchi, L., & Fontana, G. (2008). The Effectiveness of Airborne LiDAR Data in the Recognition of Channel-Bed Morphology. *Catena*, 73(3), 249-260. doi:10.1016/j.catena.2007.11.001
- Chirico, P. (2004). Surface-to-Surface Comparison of NED, SRTM, Digital Photogrammetry and LIDAR Derived Digital Elevation Models: Implications for Geological Mapping, and Geographic Analysis. *Paper Presented at USGS GIS Conference, March 1-5*. Denver, CO.
- Committee on Floodplain Mapping Technologies, N. R. (2007). *Elevation Data for Floodplain Mapping*. Washington, DC: The National Academies Press.
- CTDEEP. (2012). *Fenton River Watershed TMDL*. State of Connecticut, Department of Energy & Environmental Protection, 30.

- Deerfield River Watershed Association. (2014). *Science in the Watershed: Hydrology*. Retrieved from Deerfield River Watershed Association: <http://deerfieldriver.org/>
- Dewberry. (2011). *Project Report for the U.S. Corps of Engineers High Resolution LiDAR Data Acquisition & Processing for Portions of Connecticut*. United States Department of Agriculture, Prepared for the Natural Resources Conservation Service. United States Army Corps of Engineers.
- Dilts, T. (2011, 10 10). Polygon to Centerline. ArcGIS Resource Center. Retrieved from <http://www.arcgis.com>
- Dunne, T., & Leopold, L. (1978). *Water in Environmental Planning*. New York, NY: W.H. Freeman and Company.
- Environment Agency. (2003). *River Habitat Survey in Britain and Ireland: Field Survey Guidance Manual*. Bristol.
- ESRI. (2014, April 15). ArcGIS Desktop: Release 10.2.2. Redlands, California: ESRI.
- European Commission. (2000, October 23). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000: Establishing a Framework for Community Action in the Field of Water Policy. *Official Journal of the European Communities*, L 327, 1-73.
- Faux, R. N., Buffington, J. M., Whitley, M. G., Lanigan, S. H., & Roper, B. B. (2009). Use of Airborne Near-Infrared LiDAR for Determining Channel Cross-Section Characteristics and Monitoring Aquatic Habitat in Pacific Northwest Rivers: A Preliminary Analysis. *PNAMP Special Publication: Remote Sensing Applications for Aquatic Resource Monitoring* (pp. 43-60). Cook, WA: Pacific Northwest Aquatic Monitoring Partnership.
- Ferreira, M. (2014, December 5). Transect Tool. GIS 4 Geomorphology. Retrieved from <http://gis4geomorphology.com>
- Hohenthal, J., Alho, P., Hyypä, J., & Hyypä, H. (2011). Laser Scanning Applications in Fluvial Studies. *Progress in Physical Geography*, 35(6), 782-809. doi:10.1177/0309133311414605
- Johansen, K., Tiede, D., Blaschke, T., Arroyo, L. A., & Phinn, S. (2011). Automatic Geographic Object Based Mapping of Streambed and Riparian Zone Extent from LiDAR Data in a Temperate Rural Urban Environment, Australia. *Remote Sensing*, 2011(3), 1139-1156. doi:10.3390/rs3061139

- Kasprak, A., Magilligan, F. J., Nislow, K. H., & Snyder, N. P. (2012). A LiDAR-Derived Evaluation of Watershed-Scale Large Woody Debris Sources and Recruitment Mechanisms: Coastal Maine, USA. *River Research and Applications*, 28(9), 1462-1476. doi:10.1002/rra.1532
- Lamont, E., Farrell, J., Walker, S., & Rosa, D. (2013). *Fluvial Geomorphic Assessment of the Fenton River - Chaffeeville Silk Mill Section - Mansfield, CT*. Prepared for Geography 5310, University of Connecticut.
- Lane, E. W. (1954). *The Importance of Fluvial Morphology in Hydraulic Engineering*. Bureau of Reclamation, Engineering Laboratories Division Hydraulic Laboratory Branch. Denver, CO: United States Department of the Interior.
- McDonough, C., Mabee, S. B., & Marcus, M. (2013a). Fluvial Erosion Hazard Map of Clesson Brook: Hawley/Buckland Town Line to Cross Street, Buckland, Franklin County, Massachusetts. *Massachusetts Geological Survey, Fluvial Erosion Hazard Map FEH-13-01. Scale 1:7000, 1 Sheet and Digital Product*.
- McDonough, C., Mabee, S. B., & Marcus, M. (2013b). Fluvial Erosion Hazard Map of Pelham Brook: Pelham Lake to the Charlemont Town Line, Rowe, Franklin County, Massachusetts. *Massachusetts Geological Survey, Fluvial Erosion Hazard Map FEH-13-02. Scale 1:7000, 1 Sheet and Digital Product*.
- McDonough, C., Mabee, S. B., & Marcus, M. (2013c). Fluvial Erosion Hazard Map of the East Branch of the North River: Vermont Border to the Shulburne Town Line, Colrain, Franklin County, Massachusetts. *Massachusetts Geological Survey, Fluvial Erosion Hazard Map FEH-13-03. Scale 1:8000, 2 Sheets and Digital Product*.
- McDonough, C., Mabee, S. B., & Marcus, M. (2013d). Fluvial Erosion Hazard Map of the Green River: Leyden Town Line to the Deerfield River, Greenfield, Franklin County, Massachusetts. *Massachusetts Geological Survey, Fluvial Erosion Hazard Map FEH-13-04. Scale 1:7000, 2 Sheets and Digital Product*.
- MEEA. (2004). *Deerfield River Watershed Assessment Report: 2004-2008*. The Commonwealth of Massachusetts, Executive Office of Environmental Affairs. Boston, MA: Executive Office of Environmental Affairs.
- Montgomery, D. R., & Buffington, J. M. (1998). Channel Processes, Classification, and Response. In R. J. Naiman, & R. Bilby (Eds.), *River Ecology and Management* (pp. 13-42). New York, New York: Springer-Verlag New York, Inc.
- Northrop Grumman Corporation. (2012). *Project Report FEMA Region 1 FY12 LiDAR Task Order - Hudson-Hoosic-Deerfield NY*. Project Report, Prepared for USGS and FEMA.

- O'Callaghan, J. F., & Mark, D. M. (1984, December). The Extraction of Drainage Networks from Digital Elevation Data. *Computer Vision, Graphics, and Image Processing*, 28(3), 323-344. doi:10.1016/S0734-189X(84)80011-0
- Osborn, K., List, J., Dean, G., Crowe, J., Merrill, G., Constance, E., . . . Kosovich, J. (2001). Chapter 4: National Digital Elevation Program (NDEP). In D. Maune, *Digital Elevation Model Technologies and Applications: The DEM Users Manual* (pp. 83-120). Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing.
- Pivarnik, A., Nicoulin, A., & MicCusker, M. (2013). *Fluvial Geomorphic Assessment of the Fenton River, Gurleyville Road Reach*. Prepared for Geography 531, University of Connecticut.
- Raven, P. J., Holmes, N. T., Charrier, P., Dawson, F. H., Naura, M., & Boon, P. J. (2002, July/August). Towards a Harmonized Approach for Hydromorphological Assessment of Rivers in Europe: a Qualitative Comparison of Three Survey Methods. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 12(4), 405-424. doi:10.1002/aqc.536
- Rinaldi, M., Surian, N., Comiti, F., & Bussetini, M. A. (2013, January). A Method for the Assessment and Analysis of the Hydromorphological Condition of Italian Streams: The Morphological Quality Index (MQI). *Geomorphology*, 180-181(1), 96-108. doi:10.1016/j.geomorph.2012.09.009
- Rosgen, D. L. (1994). A Classification of Natural Rivers. *Catena*, 22, 169-199.
- Rosgen, D. L. (1996). *Applied River Morphology*. Pagosa Springs, CO: Wildland Hydrology.
- Sinha, J. (2000, February 7). *Creating DEMs using TOPOGRID*. Retrieved August 2013, from http://ibis.geog.ubc.ca/courses/geog570/talks_2000/topogrid.html
- Snyder, N. P. (2009, February). Studying Stream Morphology With Airborne Laser Elevation Data. *Eos, Transactions American Geophysical Union*, 90(6), 45-46. doi:10.1029/2009EO060001
- USGS. (2013, February 1). Mansfield; Willington. *Connecticut 2012 Multispectral Orthophotography by Town*. Storrs, Connecticut: CT Environmental Conditions Online (UConn and CT DEEP). Retrieved from CT.gov: <http://www.ct.gov/deep/site/default.asp>
- USGS. (n.d.). <http://nationalmap.gov>. Retrieved August 2013, from The National Map.

Vocal Ferencevic, M., & Ashmore, P. (2012). Creating and Evaluating Digital Elevation Model-Based Stream-Power Map as a Stream Assessment Tool. *River Research and Application*, 28, 1394-1416. doi:10.1002/rra.1523