ABSTRACT

GARCIA, VALERIE COVER. Using GIS and LIDAR to Map Headwaters Stream Networks in the Piedmont Ecoregion of North Carolina (Under the direction of Heather Cheshire.)

A large percentage of nonpoint source pollution found in our Nation’s waterbodies is suspected to occur through first- and second-order (headwaters) streams. Such streams drain a much greater proportion of watershed area and have a much greater length of riparian zone interaction with the land than the higher-order streams typically studied for nonpoint source water quality problems. The State of North Carolina and the U.S. EPA are interested in examining the contribution of lower-order streams to the overall nonpoint source pollution problem; however, the mapping of first- and second-order streams is extremely poor. The recent availability of fine resolution Light Detection and Ranging (LIDAR) data for portions of the State of North Carolina provides the opportunity for developing improved methods of mapping lower-order streams using Geographic Information System (GIS) approaches. In this study, I investigated the state-of-science for mapping topography and extracting headwaters stream networks using LIDAR data and GIS approaches. I applied these techniques to map headwaters streams at a study site in the Piedmont Ecoregion of North Carolina. I found that LIDAR produced more accurate elevation maps (elevation accuracy within 1.2’) than currently available maps, such as the USGS 7.5 minute Digital Elevation Models (elevation accuracy within 49’). The Triangulated Irregular Network (TIN) produced the best topographic maps, but the Digital Elevation Model (DEM) was better for automatically extracting headwaters streams. The best headwaters stream maps were derived by using a hydro-enforced TIN for generating the base DEM, and extracting the stream network from this base DEM using ArcHydro and the AGREE algorithm. These improved headwaters stream maps will enable decision-makers to assess and mitigate nonpoint source water quality problems.
USING GIS AND LIDAR TO MAP HEADWATERS STREAM NETWORKS IN THE PIEDMONT ECOREGION OF NORTH CAROLINA

By

VALERIE COVER GARCIA

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

FORESTRY DEPARTMENT

Raleigh
2004

APPROVED BY:

James D. Gregory, Ph.D.
George Hess, Ph.D.

Heather M. Cheshire, Ph.D.
BIOGRAPHY

Valerie Garcia received her Bachelor of Science in Information Management from the University of San Francisco in 1995. She is a mother of four children, and has worked for the Federal Government for over fifteen years. She is currently an Assistant Lab Director for the Office of Research and Development within the U.S. Environmental Protection Agency.
ACKNOWLEDGEMENTS

As with any major endeavor, my thesis is the result of the support and assistance of many individuals. My advisors patiently guided me with strong, but gently worded advice. My supervisors and co-workers juggled duties for me, provided encouragement, and continually prompted me to stay the course. Most of all, my husband and children endured much time away from me, missed events, and a somewhat harried wife and mother, yet they always supported me with love and enthusiasm. To all of you I say “thank you” from the deepest part of my heart.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td><strong>PART I: REVIEW OF HEADWATERS STREAM MAPPING TECHNIQUES USING</strong></td>
<td></td>
</tr>
<tr>
<td><strong>GEOGRAPHICAL INFORMATION SYSTEM METHODS AND LIDAR DATA</strong></td>
<td></td>
</tr>
<tr>
<td>1.0 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2.0 BACKGROUND</td>
<td>3</td>
</tr>
<tr>
<td>3.0 TOPOGRAPHIC MAPS</td>
<td>5</td>
</tr>
<tr>
<td>3.1 Model Descriptions</td>
<td>5</td>
</tr>
<tr>
<td>3.2 Storage and Computational Differences</td>
<td>7</td>
</tr>
<tr>
<td>3.3 Artifacts</td>
<td>8</td>
</tr>
<tr>
<td>3.4 Hydro-Enforcement</td>
<td>9</td>
</tr>
<tr>
<td>3.5 Interpolation</td>
<td>10</td>
</tr>
<tr>
<td>3.6 Summary of Topographic Mapping Differences</td>
<td>15</td>
</tr>
<tr>
<td>4.0 EXTRACTION OF HEADWATERS STREAM NETWORKS</td>
<td>17</td>
</tr>
<tr>
<td>4.1 Stream Extraction Methods</td>
<td>23</td>
</tr>
<tr>
<td>4.2 Discussion of Stream Extraction Methods</td>
<td>27</td>
</tr>
<tr>
<td>5.0 CONCLUSIONS AND RECOMMENDATIONS</td>
<td>28</td>
</tr>
<tr>
<td>6.0 REFERENCES CITED</td>
<td>33</td>
</tr>
</tbody>
</table>
# PART II: AUTOMATED TECHNIQUES USING GIS AND LIDAR TO MAP HEADWATERS STREAM NETWORKS IN THE PIEDMONT ECOREGION OF NORTH CAROLINA

## 1.0 INTRODUCTION

## 2.0 BACKGROUND

## 3.0 OBJECTIVES

## 4.0 METHODS

<table>
<thead>
<tr>
<th>Objective</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Evaluation of LIDAR Surface Elevation Points</td>
</tr>
<tr>
<td>2</td>
<td>Comparison of Topographic Maps Generated Using Various Interpolation Methods</td>
</tr>
<tr>
<td>3</td>
<td>Effect of Scaling on Accuracy of Topographic Maps and Extracted Stream Networks</td>
</tr>
<tr>
<td>4</td>
<td>Assessment of Hydro-Enforcement on Extracting Stream Networks</td>
</tr>
</tbody>
</table>

## 5.0 RESULTS AND DISCUSSIONS

<table>
<thead>
<tr>
<th>Objective</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Evaluation of LIDAR Surface Elevation Points</td>
</tr>
<tr>
<td>2</td>
<td>Comparison of Topographic Maps Generated Using Various Interpolation Methods</td>
</tr>
<tr>
<td>3</td>
<td>Effect of Scaling on Accuracy of Topographic Maps and Extracted Stream Networks</td>
</tr>
<tr>
<td>4</td>
<td>Assessment of Hydro-Enforcement on Extracting Stream Networks</td>
</tr>
</tbody>
</table>

## 6.0 SUMMARY AND CONCLUSIONS

## 7.0 REFERENCES CITED
## List of Tables

### PART I

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.1</td>
<td>Stream Extraction—Standard DEM Approach</td>
<td>18</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Stream Extraction—Standard TIN Approach</td>
<td>19</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Algorithms Used in Watershed Analysis</td>
<td>21</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Comparison of DEMs and TINs for use with LIDAR Data to Extract Headwaters Stream Networks</td>
<td>30</td>
</tr>
</tbody>
</table>

### PART II

| Table 4.2.1 | Interpolations Methods and Settings                                          | 49   |
| Table 4.4.1 | Interpolation and Hydro-Enforcement Method Used for Extracting Stream Networks | 55   |
| Table 5.2.1 | Comparison of Interpolated Elevation Values to Survey Measures               | 57   |
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1</td>
<td>Methodology Flow Chart: Develop Survey Plan</td>
<td>70</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Methodology Flow Chart: Compare LIDAR and Interpolation Methods</td>
<td>71</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Methodology Flow Chart: Evaluate Scaling</td>
<td>72</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Methodology Flow Chart: Effect of Hydroenforcement on Stream Extraction</td>
<td>73</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Study Area</td>
<td>74</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Close-up of Study Catchment</td>
<td>75</td>
</tr>
<tr>
<td>Figure 5.1.1</td>
<td>Comparison of Survey and LIDAR Elevation</td>
<td>76</td>
</tr>
<tr>
<td>Figure 5.2.1</td>
<td>Comparison of Survey, Spline and Natural Neighbor</td>
<td>77</td>
</tr>
<tr>
<td>Figure 5.2.2</td>
<td>Comparison of Survey, Kriging and IDW</td>
<td>77</td>
</tr>
<tr>
<td>Figure 5.3.1</td>
<td>Comparison of Survey, 10 ft. and 20 ft. resolution DEMs</td>
<td>78</td>
</tr>
<tr>
<td>Figure 5.3.2</td>
<td>Comparison of Survey, 60 ft. and 90 ft. resolution DEMs</td>
<td>79</td>
</tr>
<tr>
<td>Figure 5.3.3</td>
<td>Comparison of Stream Drainages--10 and 20 ft.</td>
<td>80</td>
</tr>
<tr>
<td>Figure 5.3.4</td>
<td>Comparison of Stream Drainages--60 ft.</td>
<td>81</td>
</tr>
<tr>
<td>Figure 5.3.5</td>
<td>Comparison of Breakline Data</td>
<td>82</td>
</tr>
<tr>
<td>Figure 5.3.6</td>
<td>Comparison NHD to Breakline Data</td>
<td>83</td>
</tr>
<tr>
<td>Figure 5.3.7</td>
<td>Hydroenforced TIN</td>
<td>84</td>
</tr>
<tr>
<td>Figure 5.3.8</td>
<td>Stream Network Extracted from enforced and unenforced TIN</td>
<td>85</td>
</tr>
<tr>
<td>Figure 5.3.9</td>
<td>Effect of using AGREE to Extract Stream Networks</td>
<td>86</td>
</tr>
<tr>
<td>Figure 5.3.10</td>
<td>“Best” Modeled Drainage</td>
<td>87</td>
</tr>
<tr>
<td>Figure 5.3.11</td>
<td>Comparison of modeled vs. manual--study headwater stream</td>
<td>88</td>
</tr>
</tbody>
</table>
1.0 Introduction

The Strahler (1952) stream ordering approach assigns the source of a stream as first order, and the confluence of two or more first-order streams as second order. This assignment of higher orders continues until the entire stream network has been numbered. Headwaters stream networks are defined as Strahler first- and second-order streams in a watershed stream network, regardless of the size of the watershed or location within the watershed that the streams occur (Strahler, 1952). Headwaters stream networks serve as a critical hydrologic link between the surrounding landscape and larger, connecting stream outflows within a watershed. These networks drain extensive surface areas of land within the watershed that are not directly in contact with higher order stream channels. As changes in the adjacent land use and land cover occur from human and natural pressures, flows occurring in these lower order streams export diffuse pollution that has been aggregated from throughout the watershed (Meyer and Wallace, 2001; Peterson, et al., 2001). The origin of poor or degraded water quality appearing in higher order streams is difficult to quantify, because the lower order streams transport materials differentially from unidentified non-point sources within the entire watershed.

Accurate stream maps are needed to investigate the characteristics, magnitude and mitigation of non-point source pollution at headwaters streams. Local, state, and federal agencies have relied on US Geological Survey 1:24,000 scale topographic maps for information on stream networks for planning, management, and regulatory programs related to streams. However, recent research in North Carolina has demonstrated the relatively low
accuracy of US Geological Survey (USGS) 1:24,000 scale maps in depicting the presence and location of low order, headwaters streams (Darling, et al. 2002; Gregory, et al., In Review, 2004). In terms of stream length, the USGS topographic maps, on average, overestimate the presence of headwaters streams in the Coastal Plain by 31%, underestimate the presence of headwaters streams in the Piedmont by 25%, and in the Mountains, by up to 54% (Gregory, et al., In Review, 2004).

Currently, the very low accuracy of topographic maps in depicting first- and second-order streams places an inordinate burden on both regulatory agencies and the regulated community to conduct joint field checks of stream locations. Improved headwaters streams maps would be useful to a wide variety of state agencies in planning activities, such as, implementing road construction projects and performing industrial recruitment efforts. Local governments would benefit from the maps in their regional planning efforts and application of riparian buffer regulations and ordinances. Thus, a new automated mapping approach for identifying headwaters stream networks would contribute substantially to state and regional planning and regulatory communities, as well as local regulated communities.

I reviewed the state of the science for using GIS approaches for mapping topography and extracting headwaters stream networks. The focus of the investigation was on the unique aspects of using Light Detection and Ranging (LIDAR) data to produce maps appropriate for the headwaters stream scale. In this review, I: (1) identified the unique aspects of using LIDAR bare-earth elevation points to produce stream drainage networks; (2) compared Digital Elevation Models and Triangulated Irregular Networks for mapping topography; (3) compared various modeling approaches used to map stream networks; and (4) recommended which approaches and/or combination of approaches are appropriate for mapping headwaters
2.0 Background

Automated stream mapping approaches require accurate topographic maps, as gravity is the most prominent factor in the drainage of water from the landscape (Maune et al., 2001; Theobald and Goodchild, 1990). The most widely available topographic maps typically used for automated stream extraction are produced and maintained by the USGS’ National Digital Elevation Program (NDEP). While these maps are being updated, the source data used to produce the bulk of these maps were derived from the digitization of contours (Digital Line Graphs) drawn on USGS 1:24,000 topographic maps, with some of the oldest maps produced with orthophotogrammetric methods (Osborn, et al., 2001). The resulting 30m or less point-spaced Digital Elevation Maps are not meant to support the resolution necessary for extracting headwaters streams (Osborn, et al., 2001).

The advent of Light Detection and Ranging (LIDAR) technology, however, has presented new opportunities in producing high-density, high-accuracy elevation data as a source for accurate topographic maps (Renslow, 2001; Maune et al., 2001; and Band, 1993). Airborne LIDAR consists of sensors that emit 5,000 - 50,000 laser pulses per second in a scanning array. The sensors measure the round trip time of the transmitted laser pulse that is reflected from a target. The sensor receives up to five returns for each emitted laser pulse. For example, the first return records the reflection from the first obstacle encountered by the pulse (e.g., canopy top), the second return represents the next obstacle encountered (e.g., bush cover), and the third return represents the final surface (e.g., ground or bare earth). Airborne LIDAR consists of a laser sensor installed on an airplane specially equipped with
global positioning system (GPS) and inertial measuring unit (IMU) technology. These technologies, combined with ground GPS measurements, continuously measure the three dimensional position and orientation of each laser pulse, resulting in highly accurate horizontal and vertical measures (NCCTS, 2003; Thompson and Maune et al., 2001; Renslow, 2001).

One benefit of using LIDAR rather than conventional methods (e.g., stereo photogrammetry, contour digitization) is the production of more accurate and densely spaced elevation points. This increase in accurate data has enhanced the quality, definition and resolution of available topographic maps (Osborn, et al., 2001, Fowler, 2001). LIDAR also results in more accurate elevation maps in vegetated areas, because LIDAR needs only a single pulse to reach the ground whereas photogrammetry requires different lines of sight to see the same points on the ground (NCCTS, 2003; Thompson and Maune, 2001; Renslow, 2001). The disadvantages of using LIDAR are: (1) LIDAR results in artifacts—erroneous bare-earth elevation determinations caused by returns from features such as, bridges, culverts and dams; (2) LIDAR pulses are fully absorbed by water, rendering water returns unreliable; (3) densely vegetated areas (e.g., forest, grass, shrub) can produce false bare-earth elevation data; and (4) post-processing removes questionable returns resulting in irregularly spaced points and sometimes large areas with no data (Renslow, 2001; Spinney, 2001; NCCTS, 2003; Thompson and Maune, 2001).

There are two main processing steps used to develop stream network maps: (1) the development of accurate topographic maps; and (2) the production of representative stream networks from these topographic maps. The strengths and weaknesses of using LIDAR data to develop stream maps become important when considering these two major processes. The
fine-resolution necessary for mapping of headwaters streams, and the processes that are relevant at this scale are also unique. For example, physical processes that are “averaged-out” over regional or watershed scales (e.g., diffuse flow) are relevant at the headwaters stream scale. In the following sections, I investigated the state-of-science for mapping headwaters stream networks, focusing on the issues pertaining to these processing steps.

3.0 Topographic Maps

Digital Elevation Models (DEM) and Triangulated Irregular Networks (TINs) are the most widely used spatial model types for developing topographic maps. DEM is used here to describe a topographic map based on a raster or grid structure, whereas, TIN is used to describe a topographic map based on a vector structure (points, lines and polygons). LIDAR data can be used to produce topographic maps using either model type. Both the DEM and TIN modeling approaches offer strengths and weaknesses in the basic premises underlying the models. The following section investigates these differences, including differences in the storage and data structures, computational efficiencies, errors introduced by the models, and the role of hydro-enforcement and interpolation.

3.1 Model Descriptions

DEM is used in this paper to describe a uniformly spaced, 3-dimensional cartographic representation (x,y, and z) in a grid or raster\(^1\) format (Figure 3). This structure is created by interpolating sample points to derive a value for each uniformly-spaced point or area.

\(^1\)As used by some authors, grid assumes the assignment of the data value to the pixel center, whereas raster assumes the assignment of the value to the pixel area. In this paper,
Triangulated Irregular Network (TIN) is used to describe the formulation of non-overlapping triangles from irregularly spaced x, y, and z points. TINs are vector models, having attributes of slope, aspect, and area, with the three vertices having elevation attributes, and the three edges having slope and direction attributes (Figure 2). Triangle vertices are connected in Delaunay triangulation scheme producing the preferred ~ 60 degree angles; the formation of equilateral triangles ensures that any point on the surface is as close as possible to a vertex (Poiker, 1990; Maune et al., 2001). Newer TIN models incorporate a mathematical function that ensures that the slope changes continuously, rather than abruptly, at the triangle edges (Poiker, 1990; Theobald and Goodchild, 1990).

Figure 1 Digital Elevation Map. Elevation values are assigned to center point in grid structure (right); resulting topographic map for Falls Lake, North Carolina (left)

---

grid and raster are used synonymously to describe a regularly spaced grid structure.
3.2 Storage and Computational Differences

LIDAR produces large datasets in comparison with standard data collection techniques, (e.g., deriving points from photogrammetry, existing DEMs, or survey). The storage and computational challenges presented by these datasets are further confounded because of the fine resolution required for extracting headwaters stream networks. While DEMs offer advantages in storage and computational efficiency, it is important to note that the number of grid cells increases four-fold for each doubling of the horizontal DEM resolution (Band, 1993; Nelson, et al., 1999; Maune et al., 2001). Nevertheless, the storage and manipulation of data in grid structures is straightforward, as the horizontal location is implicit in the order of the grid (Kumler, 1990).

Conversely, there is no implicit topology within the TIN structure, requiring that connectivity between nodes be defined explicitly. This, coupled with the need to store the location of every vertex in x-, y- and z-dimensions, requires more storage than grid models²

---

²While grid structures are inefficient at storing redundant information (e.g., elevation grid values for homogenous terrain) (Band, 1993), compression techniques, such as run-length encoding are available to minimize storage space when adjacent cells have the same value (Maune, et al., 2001).
Once the TIN topology is defined, however, processing is more efficient for functions such as spatial searching, because these relationships do not need to be computed with each application as is done with DEMs (Maune et al., 2001).

DEM models are the most common form of terrain models currently in use and, as a result, more analysis tools are available for this structure. Some math functions, such as map algebra (e.g., computing grid values using one or more layers), are more easily performed using DEMs, while trigonometrically based values, such as slope, aspect, and area, are more easily computed using the TIN (Poiker, 1990; Theopald and Goodchild, 1990; Nelson, et al., 1999). This advantage of the TIN structure becomes particularly significant when the topographic map is the basis for hydrologic modeling. Theobald and Goodchild (1990) make the convincing argument that aspect is the major criterion to assess terrain accuracy for the purpose of hydrological modeling, because aspect is the primary determinant of local flow direction. They found that DEMs are most likely to contain errors in slope and aspect estimation where there is rapid elevation change or where the slope is not continuous, such as channels, and ridges.

### 3.3 Artifacts

Another challenge presented by both the DEM and TIN terrain models is the occurrence of artifacts, or errors generated by the modeling techniques that cause false high or low points in the modeled terrain. These unnatural undulations can interfere with the simulation of water flow. DEMs often produce spurious depressions or pits due to noise in the elevation data, making it difficult to tell if the pits are natural and should be there or if
they are errors introduced by the model (DeVantier, B. and A. Feldman, 1993; Maune et al., 2001). TINs avoid this to some extent as the movement of water follows the slope of a plane or flows down the edge between two triangles (DeVantier, B. and A. Feldman, 1993; Nelson, et al., 1999; Maune et al., 2001). TINs, however, are prone to the creation of triangles that erroneously cross over and block natural drainage paths (Theopald and Goodchild, 1990; Mitas and Mitasova, 1999; Nelson, et al., 1999). The use of LIDAR surface elevation points confounds the problem of drainage interruptions, as returns from surfaces such as, culverts, bridges, and boulders create terrain features that artificially block the natural flow of water (Renslow, 2001; Thompson and Maune, 2001).

3.4 Hydro-enforcement

To solve the problem of drainage obstructions, some DEM and TIN software applications allow for hydro-enforcement, the incorporation of known stream or ridge center-lines (commonly referred to as breaklines) into the modeled terrain to remove undulations in the surface and ensure the downstream drainage of water. Whether using the DEM or TIN modeling approach, several authors emphasize the need to supplement elevation data with breakline data (Theopald and Goodchild, 1990; DeVantier, B. and A. Feldman, 1993; Maune et al., 2001; Thompson and Maune, 2001; NCCTS, 2003). It is important to note that the best source for deriving breaklines is through digitizing orthophotographs; use of DEMs or TINs to produce breakline data has not been successful (Maune et al., 2001; Thompson and Maune, 2001; NCCTS, 2003). The need to construct breakline files adds significantly to the resources required to produce accurate topographic models.

In DEMs, stream or ridge center-lines are represented as a series of adjacent grids and
are overlain onto the grid structure after construction, or incorporated as part of the grid development using an iterative approach. For TINs, the breaklines are incorporated into the triangulation scheme, so that the center-line becomes the edge of the bounding triangles. In both cases, the known stream or ridge center-lines become the forced low- or high-point in the topography. Hydro-enforcement works particularly well in combination with the TIN, as the center-lines form the natural TIN edges during the formulation of the topography (Poiker, 1990; Theobald and Goodchild, 1990; Maune et al., 2001).

The fine-resolution necessary for extracting headwaters streams presents two challenges with regard to hydro-enforcement. First, the resolution of orthophotographs typically used for deriving breaklines does not allow for the identification of headwaters streams. While breakline data are valuable for ensuring downstream drainage for higher order streams where breakline data are available, the problem of modeled artifacts interrupting drainage is still a problem for extracting headwaters streams. Second, the relatively accurate and fine resolution offered by topographic maps constructed using LIDAR data makes it difficult to discern true surface variations from artifacts—particularly for headwaters stream terrain. The characteristics of headwaters streams (e.g., head cuts, variability of distinct channel formation) presents challenges in determining if an obstruction to drainage is due to natural terrain variation, or if it is an artifact introduced by the modeling technique.

3.5 Interpolation

Interpolation is used to estimate values for the regularly spaced DEM from the irregularly spaced LIDAR data. The various interpolation methods used to estimate
unsampled elevation points have strengths and weaknesses that can impact the accuracy of
that, “All the major difficulties with computer interpolation are caused by insufficient data
and observational error”. More accurate data equates to more accurate interpolation--
regardless of the interpolation method used. With regard to LIDAR data, the density of
sample points is greatly increased over standard methods. LIDAR data, however, results in
voids, or areas with no point coverage, due to errors introduced by the technology, creating
unique interpolation challenges. Even in the case of TINs where interpolation is not used to
create the modeled surface, Spinney (2001) found that the use of LIDAR data as a source for
creating a TIN resulted in large, unnatural triangles, caused by the large data voids inherent
with this technology. Thus, the interpolation of estimated values from known values is one
of the most important components of producing accurate topographic maps.

The interpolation of measured points is used to: (1) convert irregularly spaced points to
a grid structure; (2) locate a query point (other than the triangle vertices) on a TIN; and (3)
resample to different resolutions. For purposes of modeling terrain, there are two main types
interpolation methods account for the larger of the two categories and includes
methods based directly on surrounding measured values and/or mathematical formulas
applied to those values. Kriging is a geostatistical approach that relates the strength of
similarity between measured samples accounting for distance and direction, and is the only
probabilistic model discussed here.

Inverse Distance Weighted (IDW): IDW determines new cell values using a linearly
weighted combination of values from nearby points. Weights are usually inversely
proportional to a power of distance. While this method is appropriate when a linear relationship exists with proximity to a sample point, this assumption does not hold true for terrain data and results in a dimpled effect throughout the simulated terrain (Maune et al., 2001; Mitas and Mitasova, 1999).

**Natural Neighbor:** The Natural Neighbor interpolation method (also called area stealing), uses an area based weighting scheme to incorporate distance and proximity. This method is often used in determining an unsampled point in a TIN. The Thiessen polygon, boundaries created from assigning locations to the nearest vertex (Poiker, 1990), is determined for the queried unsampled point, and overlain over the Thiessen polygons resulting from the sampled points. The area of the overlap is used to weight the influence of the local sampled points, resulting in a surface with smoothly changing gradients (Watson, 1992, Maune et al., 2001). The natural neighbor interpolation method performs well when used with second-order derivatives to produce slopes that change continuously, rather than abruptly, at the triangle edges. (Poiker, 1990; Theobald and Goodchild, 1990, Maune, et al., 2001).

**Spline:** The spline technique is often described as bending a rubber sheet through a set of points. The method minimizes the overall curvature of the surface, resulting in a smoothed surface that passes through the sample data points (Maune et al., 2001). Mitas and Mitasova (1999) describe a method called Regularized Spline with Tension (RST). RST allows for the tuning of the surface tension, modifying the rubber sheeting effect from a stiff plate into an elastic membrane. Parameters such as tension and smoothing weights can be adjusted to prevent artificial features such as waves along contours, and artificial peaks, pits
or overshoots. Parameters can be selected empirically, based on knowledge of the modeled phenomenon, or automatically by minimizing the predictive error estimated by a cross-validation procedure (Mitasova, et. al, 1995). The tension parameter can be used to enable modeling of anisotropy (showing different properties in different directions). For example, this parameter was used to improve predictions of local precipitation in mountainous terrains (Hofierka et al., 2002).

**Kriging:** Kriging weights the surrounding measured values to derive a prediction for each location. Weights are based on a combination of the distance between the measured points, the prediction location, and the strength of the overall correlation among the measured points. A semivariogram, depicting the difference in z-values, and the mapped distance and direction between coordinate pairs, is computed for the sample data points. Fitting the model is based on selecting the appropriate parameters that fit the best line through the semivariogram. The parameters of these functions are then optimized for the best fit of the experimental semivariogram (Cressie, 1993). The interpolated surface is then constructed using statistical conditions of unbiasedness and minimum variance. The main strength of kriging is in the statistical quality of predictions (e.g., unbiasedness) and the ability to predict the spatial distribution of uncertainty (Isaaks and Srivastava, 1989). Kriging is sensitive, however, to the homogeneity of the surface. Maune et al. (2001) gives the following example: “if you had a point dataset for the entire United States and fit a model, it would be much different than the model of the state of Kansas.”

**Summary of Interpolation Methods:** Data points used in interpolation are often spatially heterogeneous and sampling data can often be far from optimal. Data can contain significant noise or discontinuities and data sets can be very large, originating from a variety
of sources with variable accuracies (Mitas and Mitasova, 1999). All of these factors are important in considering the appropriate interpolation scheme. For example, in the case of using raw LIDAR data (no post-processing applied to determine bare-earth), the spline with tension approach performed well in defining terrain at a coarse scale. The spline with tension function smoothed out the “noise” associated with the multiple LIDAR returns to generate the general topography of the land. This same study also revealed that kriging estimates provided reasonable results in regions with sufficiently dense observations, but in mountainous regions with sparse data it did not reflect the impact of the topographical patterns (Hofierka, et al., 2002).

Inverse Distance Weighting is the least preferred interpolation method for modeling terrain, because terrain is not linearly related to the proximity between sample points (Maune et al., 2001; Mitas and Mitasova, 1999). Natural Neighbor, typically used to determine non-sampled points in a TIN model, works well when second-order derivatives are used to ensure continuous slope at triangle edges (Theobald and Goodchild, 1990; Nielson, 1983).

The kriging and regularized spline with tension techniques offer the two most viable options for interpolating irregularly spaced points to a grid-based structure (Maune, 2001; Mitas and Mitasova, 1999). The advantage of kriging is in the statistical basis of the resulting predictions (unbiasedness) and in the ability to predict the spatial distribution of uncertainty. Mitasova (1995) argues that kriging is best for those phenomenon with a strong random component or for estimation of statistical characteristics (uncertainty), but is less successful for applications where local geometry and smoothness are critical. The regularized spline with tension approach, however, works well when modeling physical phenomena that result from processes which minimize energy, (e.g., terrain with its balance
between gravitation force, soil cohesion and impact of climate) (Mitas and Mitasova, 1999). Both the spline and kriging interpolation techniques require greater knowledge and experience of the user because pre-selected model values can significantly affect the final results.

### 3.6 Summary of topographic mapping differences

The basic underlying premises of DEMs and TINs are substantially different; DEMs are based on a grid structure and result in uniformly-spaced estimated elevation values, and TINs are developed through triangulation, maintaining the original input data as the triangle vertices. When using LIDAR data to form a DEM, interpolation algorithms are used to derive uniformly-spaced values, resulting in the loss of precision and generalization of the terrain surface (Dodson and Li, 2000; Maune et al., 2001). Because the TIN maintains the original input values, the resulting surface better represents terrain variability and definition (Kumler, 1990; Nelson, et al., 1999; Dodson and Li, 2000; Maune et al., 2001; Wang, 2001). According to Poiker et al. (1990), and Theobald and Goodchild (1990), these advantages of the TIN structure are particularly relevant for fluvial eroded terrain, where TINs do particularly well at modeling topography.

The large data files created by TINs and the need to compute and store topology may present problems with storage and computational requirements—particularly at the fine resolution necessary for mapping headwaters streams. In addition, TINs offer fewer analysis tools and do not handle map algebra (the ability to compare one layer to one or more other layers) for incorporating the many grid-based data available (e.g., land use/cover, vegetation and soil maps).
For those areas where the LIDAR coverage is dense, the various interpolation approaches (natural neighbor, kriging, and regularized spline with tension) should produce similar results. For the TIN methodology, the natural neighbor interpolation method is often used. This option (enhanced by smoothing functions) does well where there is good LIDAR coverage, but does poorly when voids are encountered, producing large triangles with widely spaced vertices (Spinney, 2001). Thus, a combined approach may take advantage of the incorporation of the LIDAR source data and breaklines into the triangulation structure through the TIN model, yet leverage the strong interpolation capabilities offered by DEMs through conversion of the TIN to a grid.

Both kriging and spline should be considered for interpolating raw data points to a grid structure. With the denser coverage offered by LIDAR, kriging should perform well as the strong statistical representation of the terrain will produce more precise interpolated points. Kriging also offers the ability to quantify uncertainty. Application of kriging across large, heterogeneous areas should be avoided, however. For example, results may be more robust if the mountain and piedmont regions of North Carolina are processed using watershed boundaries instead of combining these areas across artificial boundaries (e.g., county lines). Regularized spline with tension also offers an intriguing solution and should be investigated as well. Through the variational approach, Mitas and Mitasova (1999) demonstrate the power of this interpolation approach and the ability to incorporate stream enforcement through numerical solutions, include additional conditions, and prescribe derivatives at given or arbitrary points.
4.0 Extraction of Headwaters Stream Networks

An accurate topographic map, constructed using a DEM or TIN structure, is the basis for the extraction of stream networks and is the starting point for drainage analysis (the term, stream extraction, is used here to indicate the modeling of flow paths, rather than the modeling of channel morphology or other hydraulic processes that could be assumed with the term, stream mapping). The complexity associated with stream extraction has resulted in the creation of several approaches that imbed algorithms as a series of steps in either a DEM, or TIN model. In this section, we discuss various models representing these approaches, including, the Environmental Systems Research Institute's (ESRI’s) ArcHydro Tools (2002), Tarboton’s terrain analysis tool (2002), Geographical Resources Analysis Support System (GRASS) (US Army CoE, 1999), and the Watershed Management System (Nelson, et al., 1994). These models were selected for review because they: (1) provide a broad range of modeling approaches; (2) cover the standard approaches currently used in mapping streams; and (3) take stream extraction methods the next step by offering more sophisticated techniques.

Investigation of the many GIS stream extraction methods currently available revealed that most models incorporate a basic set of steps or approaches that have become standard for DEM- and TIN-based stream extraction methods. These standard approaches are introduced below, followed by a more extensive discussion about how the selected models offer advances that distinguish them from these standard approaches.

Typical steps in determining drainage using a DEM-based model include the filling of spurious pits, calculating flow direction, determining flow accumulation, and defining streams based on a user-defined threshold for accumulation (Table 4.1). TIN-based models
determine drainage by classifying triangle edges as either confluent (flow converges from adjacent planes); diffluent (flow diverges into adjacent planes); or transfluent (flow is directed from one plane to another) (Theopald and Goodchild, 1990; Band, 1993). Thus, diffluent edges form ridgelines, confluent edges form channel lines, and transfluent edges represent overland flow. Apart from tracing the flow path, the other steps required for extracting stream networks is essentially the same for TINs as it is for DEMs (Table 4.2).

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydro-Enforcement</td>
<td>Breaklines are incorporated by converting vector linear drainage pattern to raster and overlaying on DEM.</td>
</tr>
<tr>
<td>Fill Sinks</td>
<td>Fill pits and depressions that may interrupt flow network by increasing the elevation of pit cells to the level of the surrounding terrain.</td>
</tr>
<tr>
<td>Flow Direction</td>
<td>Define the direction of the steepest descent for each terrain cell.</td>
</tr>
<tr>
<td>Flow Accumulation</td>
<td>Determine the number of upstream cells flowing into a given cell.</td>
</tr>
<tr>
<td>Stream Delineation</td>
<td>Classify all cells as either meeting or not meeting user specified flow accumulation threshold (e.g., point at which accumulating flow causes stream flow). Also known as surface area or contributing drainage area threshold.</td>
</tr>
<tr>
<td>Stream Segments</td>
<td>Identify sections of the stream channel that connect two successive junctions, a junction and an outlet, or a junction and a drainage divide.</td>
</tr>
<tr>
<td>Outlet Specification</td>
<td>Automatically determine outlets based on most downstream cell of stream segments.</td>
</tr>
<tr>
<td>Watershed Delineation</td>
<td>Delineate areas draining to outlets.</td>
</tr>
</tbody>
</table>
Table 4.2 Stream Extraction – Standard TIN Approach (Nelson, et al., 1999; Theobald and Goodchild, 1990)

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydro-Enforcement</td>
<td>Breaklines are incorporated directly through triangulation. Drainage lines become triangle edges.</td>
</tr>
</tbody>
</table>
| Trace Flow Path     | Flow paths are traced across triangle faces by determining the path of maximum downward gradient. Flow path based on three cases:  
|                     | 1. If adjacent triangle slopes away from common edge, flow path continues across face of triangle;  
|                     | 2. If adjacent triangle slopes toward the common edge, flow path continues along common edge;  
| Stream Segments     | Identify sections of the stream channel that connect two successive junctions, a junction and an outlet, or a junction and a drainage divide. |
| Outlet Specification| Automatically determine outlets based on most downstream cell of stream segments. |
| Watershed Delineation| Delineate areas draining to outlets. |

There are many approaches for determining flow direction, solving depressions, and extracting streams (Table 4.3). Single Flow Direction (SFD), particularly the D8 algorithm, is the standard approach used to determine drainage in DEM models and is available through most software packages. The D8 algorithm assigns flow to one of eight neighboring cells (ESRI, 1994). It does an excellent job of defining channels, but a poor job of describing drainage on hillslopes where diffuse flow predominates (Tarboton, 1997). In addition, the D8 algorithm produces "grid bias", resulting in stream drainages with a strong horizontal bias. The D infinity (D4) algorithm assigns flow to an aspect value between $1^\circ$ and $360^\circ$. 
Even when used with SFD, D4 improves upon results by smoothing grid bias (Neteler and Mitasova, 2002b; Tarboton, 1997; Tarboton and Ames, 2001).

Both D8 and D4 can be used with a Multiple Flow Directions (MFD) algorithm, allowing the assignment of flow in two or more directions. This approach more accurately reflects diffuse flow on hillslopes, but can create too much diffuse flow causing large data storage issues, and a lack of defined channels (this problem is particularly relevant for extracting headwaters streams). In addition, the MFD approach is difficult to use in calculating contributing drainage area. Contributing drainage area is calculated as the number of grid cells that drain through each grid cell multiplied by the grid cell area (Tarboton and Ames, 2001; Tarboton, 1997). This calculation is then used to determine the support threshold value--an estimate of the number of grids constituting a contributing drainage area that results in the origination of surface flow (stream origin). Because the MFD approach allows drainage flow across multiple grid cells, it confounds the standard SFD-based approach in determining contributing drainage area.

TINs offer a more simplified approach to determining flow direction because the calculation of slope and aspect (and therefore the maximum downward gradient) is more straightforward than for a DEM. Nelson, et al. (1999) explain that: "When flow moves across the face of the triangle, it is termed 'overland flow'. When it moves along the triangle edges, it is termed 'channel flow'." Thus, the problems encountered with multiple versus single flow direction, and D8 versus D4, are avoided in the TIN model. The TIN structure, however, presents problems with ensuring drainage. Because of the irregular nature of the TIN, it is more difficult to automatically identify and correct artificial obstructions to
drainage in a TIN than in the uniform, grid-based DEMs (Olivera, et al., 2002). In addition, there are no methods for automatically determining stream origination using a TIN model, and such determinations need to be done manually.

**Table 4.3 Algorithms Used in Watershed Analysis (Mitasova, personal communication)**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flow Direction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D8</td>
<td>Typically one-directional flow; finds steepest gradient and assigns cell flow in one of 8 directions</td>
<td>D8 + fill good for mapping streams, but does not work well for small hillslopes (creates straight lines for ephemerals)</td>
</tr>
<tr>
<td>D Infinity</td>
<td>Single or multiple directional flow; not constrained to 8 directions; averages aspect to obtain value from 1 - 360°</td>
<td>Helps eliminate grid bias; Tarboton uses limit of two directional flow to simulate diffuse flow, but still maintains channelization</td>
</tr>
<tr>
<td>Single Flow Direction (SFD)</td>
<td>Both D8 or D4 are single-flow examples (although D4 can also be multiple-flow)</td>
<td>GRASS (R.flow) uses D4 SFD, but depressions treated as sinks; ESRI uses D8 SFD</td>
</tr>
<tr>
<td>Multiple Flow direction (MFD); diffuse flow</td>
<td>D8 or D4 weighted using % of direction of multiple potential paths. D8 based on neighboring cells; D4 based on aspect</td>
<td>MFD is appropriate for hillslopes but can create diffusion of stream; Tarboton uses D4 MFD limited to 2 directions; GRASS (r.terraflow) uses D4 MFD</td>
</tr>
<tr>
<td>Algorithm</td>
<td>Description</td>
<td>Comments</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Solving Depressions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fill</td>
<td>Fills spurious depressions/ponds; can misrepresent topology if natural surface has pits/ponds; creates plateaus</td>
<td>D8 + fill good for mapping streams, but does not work well for small hillslopes (creates straight lines for ephemerals); creates larger flooded areas when used with diffuse flow</td>
</tr>
<tr>
<td>Shortest path</td>
<td>Keeps true elevation; looks down the path to determine flow</td>
<td>Good for radar; not used with diffuse flow; GRASS (r.watershed = D8 SFD, shortest path). Does nice job on streams, but can create straight lines on hillslopes</td>
</tr>
<tr>
<td>Carving</td>
<td>“Carves” through high points instead of filling low points</td>
<td>TOPAZ, River Tools. Good for “noisy” data such as radar or raw lidar data</td>
</tr>
<tr>
<td><strong>Stream Extraction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow Accumulation</td>
<td>Contributing drainage area; calculated as number of grid cells that drain through each grid cell x grid cell area</td>
<td>This algorithm is confounded when MFD is applied because multiple cells are delivering flow from and to multiple cells</td>
</tr>
<tr>
<td>Flow Accumulation Threshold (aka Support Area or Drainage Area Threshold)</td>
<td>Usually a subjective determination based on slope and area (although more objective approaches have been proposed) that predicts at what point accumulation initiates a stream (stream origin)</td>
<td>Montgomery and Dietrich (1992): $aS^n &gt; C$ where $n=2$ and $C=200m$; Barbera and Roth (1994): $AS^k$; A=contributing area, S=slope and k=function of the fractal dimension of single rivers and river networks; Tarboton and Ames (2001): based on weighting upslope curvature and Strahler drop order value</td>
</tr>
</tbody>
</table>
4.1 Stream Extraction Methods

In this section, we discuss selected stream extraction methods and how these methods address challenges such as, ensuring drainage, defining flow, and determining stream origin.

**ESRI ArcHydro Tools:** ArcHydro was developed by ESRI (2002) in conjunction with the University of Texas. The suite of ArcHydro tools are used in an ArcGIS environment to develop drainage networks, and relate these networks to other data layers (e.g., gauge locations, time series data) to support hydrologic and hydraulic analyses. The basic concepts behind the stream extraction method (called "terrain preprocessing" in ArcHydro) are not much different than the algorithms used in the ArcInfo Grid commands or the ArcView 3.x hydrology extension (e.g., the standard D8 algorithm is used to create a single-flow path through the catchment). ArcHydro, however, incorporates an improved DEM surface reconditioning algorithm called AGREE (Hellweger, 1997). In addition to filling spurious pits, the AGREE algorithm adjusts the elevation of an area around a stream to allow a gradual descent to the channel.

An important underlying premise associated with ArcHydro is the use of both grid and vector approaches to develop the stream network. The grid layers depicting features such as catchments and stream segments are converted to vector-based polygons and lines to take advantage of the strengths of both approaches (e.g., grid to ensure drainage and vector to define relationships between features). What sets this model apart from other models is the development of a geometric network. The geometric network is developed through a relational database that assigns permanent relationships between vector-based data layers and other point data. For example, a catchment is related to a stream through parent-child
relationships. In turn, the stream is related to a lake, bay or ocean. Points along the stream can be related to a gauge location and the time-series values associated with that gauge. This geodatabase connectivity allows for the synthesis of geospatial and temporal data, making it possible to construct time-sequenced maps of rainfall and the water conditions of the landscape (Maidment et al., 2002).

**GRASS (Geographical Resources Analysis Support System):** The open-source structure of GRASS (1999) has encouraged innovative advances in GIS modeling as scientists throughout the world continue to build upon this readily available model and source code. Innovations developed through GRASS often become GIS standards years later. GRASS includes more than 350 modules, and can be run under a variety of computing platforms, including Unix/Linux, and Windows environments. The accessibility of GRASS modules, however, is greater for Unix/Linux, and these programs seem better maintained and are more fully tested than Windows versions.

GRASS is for the sophisticated and experienced user. While some interfaces have been developed, knowledge of the command language, the algorithms and how to appropriately apply them are critical in obtaining valid results. GRASS has been used, however, by scientists to do extensive hydrologic and hydraulic modeling (Mitas and Mitasova, 1999, Mitasova et al., 1995, Mitasova and Mitas, 2001, Neteler and Mitasova, 2002). A range of algorithms offered in GRASS provide the flexibility to delineate single- or multiple-flow directions, apply various corrections for spurious pits (e.g., fill, shortest path, carving), and determine flow accumulation based on slope and velocity, as well as the commonly used aspect value. In particular, the ability to apply and combine physical-based algorithms, defining such quantities as concentration, density and velocity, or behavior of a system in
space and time, makes this tool extremely flexible. For example, most GIS models consider flow velocity to be uniform in calculating this value. GRASS, however, includes flow velocity based on topography and land cover to derive a more realistic solution (Mitasaova and Mitas, 2001).

**TauDEM (Terrain Analysis using DEMs):** Tarboton has led research in areas such as grid-based flow direction, flow accumulation, and channel and watershed delineation (Tarboton, 1997; Tarboton and Ames, 2001). Among other advances, Tarboton (1997) has developed a multiple flow direction approach that proportions drainage to accommodate dispersed flow on hillslopes, yet avoids problems typically encountered with MFD approaches. By limiting diffuse flow to two directions, the D4 algorithm allows the use of the contributing drainage area algorithms, jeopardized by other MFD approaches, and avoids the problem of too much diffusion of water flow on hillslopes.

In addition, Tarboton and Ames (2001) have addressed the difficult issue of determining the point at which stream flow originates by developing an objective technique to estimate the flow accumulation threshold (termed "support threshold value" by Tarboton). The flow accumulation threshold is an user-defined estimate of the number of grids accumulating flow that will produce stream flow. The challenge with any grid-based stream extraction technique is determining the right accumulation threshold value. Tarboton and Ames (2001) suggest an objective approach for determining this value by first considering the extent of upslope curvature that exists in the contributing drainage area through a weighting scheme. The authors then suggest an iterative approach between establishing the threshold and evaluating Strahler drop order values until the optimal relationship is achieved.
The Strahler stream drop value is calculated as the mean difference in elevation between the beginning and end of each stream segment, in each classification in a catchment. Tarboton and Ames (2001), have investigated the relationship between the drainage density of a stream network, and the mean stream drop values for first-order streams and all other streams. They found that the size of the contributing drainage area is optimal when the mean stream drop value for first-order streams is not statistically different from the mean stream drop value for all other stream orders.

**Watershed Management System (WMS):** While WMS (Nelson, et al., 1994) can now be used to perform a full-suite of hydrologic and hydraulic analyses using both TINs and DEMs, the model was developed originally to perform basin delineation and drainage analysis using TINs. While determining flow direction is simplified with the TIN, drainage problems produced by flat triangles (no drainage path), pits (no drainage outlet), and false dams (triangle edges that lie perpendicular to the stream) must be removed before doing drainage analysis (EMRL, 2002). WMS contains tools to automatically edit flat triangles and pits, but the irregular structure of the TIN complicates the process, and these tools do not correct all occurrences (EMRL, 2002). In addition, there are no tools to correct false dams. The incorporation of breaklines helps substantially with preventing the occurrence of these drainage obstructions, but thin boundary triangles can be introduced by the integration of breaklines into the triangulation process (Nelson et al., 1999). WMS Version 7.0 improves upon available methods by iteratively constructing triangulation around the breaklines to prevent the development of slivers. WMS does not provide an automated method, however, to determine stream origination (although flow can be calculated using the TIN as a mesh for a finite-element model).
4.2 Discussion of Stream Extraction Methods

TIN-based approaches handle the calculation of flow direction more easily than DEM approaches, avoiding the need to assess various flow solutions (e.g., SFD vs. MFD, D8 vs. D4). TIN-based methods, however, present difficulties in automatically finding and correcting drainage problems, particularly false dams. While the incorporation of breaklines helps to alleviate the occurrence of these problems, stream center-line information is rarely available for headwaters streams. TIN-based methods also present challenges in automatically determining the point at which a stream originates along a modeled drainage flow path.

Unlike TIN-based approaches, DEMs require the adoption of single flow solutions or coping with the complexities of multiple flow algorithms. The grid-based structure of DEMs can also produce a horizontal bias in the modeled stream network, particularly on hillslopes. This horizontal bias can be improved through the incorporation of multiple flow direction techniques that limit the number of possible flow paths (Neteler and Mitasova, 2002; Tarboton and Ames, 2001; Tarboton, 1997; DeVantier and Feldman, 1993). DEMs, also offer more options in hydrologic modeling. For example, there are several options for the automated correction of drainage and for determining stream origination.

Some of the models discussed above, such as GRASS, offer more robust approaches than others but are also more complex requiring substantial knowledge and experience from the user in order to accurately use and apply the model. Interfaces can be developed to ease this problem, but the issue of ease of use must be considered when determining the best approach for a given application. In addition, available technical support and maintenance of
the applications must be considered. For example, Tarboton himself recognizes the problem of updating, maintaining and supporting the TauDEM GIS Toolbar application.

5.0 Conclusions and Recommendations

Accurate stream maps begin with accurate topographic maps. New opportunities in producing accurate topographic maps now exist with the advent of LIDAR. LIDAR produces dense, irregularly spaced elevation points more accurately than ever before. While LIDAR presents new opportunities, characteristics of the data also present new challenges in developing accurate maps.

The basic underlying premises of DEMs and TINs are substantially different; DEMs are based on a grid structure and result in uniformly-spaced estimated elevation values, and TINs are developed through triangulation, maintaining the original input data as the triangle vertices. LIDAR data can be used with either model to produce topographic maps. These topographic maps become the basis for extracting stream networks. The differing advantages and disadvantages of using DEMs and TINs to automatically extract stream networks are outlined in Table 5.1, and result in the following critical conclusions:

LIDAR data present new challenges in automatically extracting stream networks. While the generally higher density of the LIDAR bare-earth elevation points vastly improves interpolation results, voids in the terrain data caused by factors such as absorption and/or post-processing present interpolation challenges. LIDAR data also produce artifacts that can affect the flow of water as predicted by GIS stream network algorithms, requiring hydro-enforcement (stream centerlines used to enforce drainage, commonly referred to as “breaklines”) (Renslow, 2001; Spinney, 2001; NCCTS, 2003; Thompson and Maune, 2001).
**TINs produce more precise topographic maps.** TINs maintain the LIDAR elevation points by incorporating them into the triangle vertices. Conversely, DEMs interpolate the mass elevation points, resulting in a loss of precision. Thus, TINs better represent the natural variability of the terrain and retain linear structures (e.g., stream channel lines or breaklines) better than DEMs (Kumler, 1990; Nelson, et al., 1999; Dodson and Li, 2000; Maune et al., 2001; Wang, 2001).

DEM are better for automatically extracting headwaters streams. Both the DEM and TIN will produce artifacts obstructing drainage (e.g., pits, flat areas, false dams). Because of its irregular structure, it is more difficult to automatically correct drainage problems in a TIN than in a DEM (Olivera, 2002, EMRL, 2002). In addition, TINs do not offer an automated way to determine stream origination (EMRL, 2002).

The interpolation method and the resolution used to generate the topographic map can impact the accuracy of the map. The various available interpolation methods produce different results, particularly in areas where there are sparse data. Some of the more complex interpolation methods (e.g., kriging) require more knowledge by the user and are computationally demanding, but are expected to perform better in modeling terrain. Related to this issue is the question of scale: at what resolution (i.e., cell size) is the headwaters stream channel definition lost? The resolution drives the storage and computational demands of the interpolation method and can impact the selection of which interpolation method can be used (Band, 1993; Nelson, et al., 1999; Maune et al., 2001). The resolution also impacts the amount of processing time required for the stream extraction programs.

**Hydro-enforcement of the DEM enhances the accuracy of the extracted stream networks.** Several authors indicate that hydro-enforcement of the TIN or DEM (e.g.,
incorporating the stream centerlines into the TIN triangle edges or overlaying them onto the DEM) contribute significantly to improving stream extraction results (Theopald and Goodchild, 1990; DeVantier, B. and A. Feldman, 1993; Maune et al., 2001; Thompson and Maune, 2001; NCCTS, 2003). Breakline data, however, are expensive to develop, as they must be derived from orthophotographs.

The physical processes relevant at headwaters stream scales are different than watershed scales typically modeled. Physical characteristics, such as overland flow, that are “averaged-out” at coarse scales (e.g., basin or watershed) become critical at fine scales (e.g., headwaters stream catchments). Thus, accurately modeling hydrologic characteristics such as stream origin and hillslope flow may require more sophisticated methods (e.g., diffuse flow) than the standard algorithms commonly used (Tarboton and Ames, 2001; Tarboton, 1997).

As a result of these conclusions, we recommend the following steps to further investigate the automated stream mapping techniques using LIDAR data:

1. Generate a TIN from LIDAR surface elevation data for comparing the accuracy of DEMs.

2. Compare the accuracy of DEMs generated using the Nearest Neighbor (TIN-to-grid), IDW, Spline and Kriging interpolation methods.

3. Evaluate the impact of scaling by generating topographic and stream network maps at various scales and comparing the accuracy of these maps.

4. Assess the effect of hydro-enforcing DEMs to generate stream maps.

5. Examine the results of more complex stream flow and origin algorithms as compared to simpler algorithms.
### Table 5.1 Comparison of DEMs and TINs for use with LIDAR Data to Extract Headwaters Stream Networks

<table>
<thead>
<tr>
<th>Criteria</th>
<th>TIN</th>
<th>DEM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Availability</strong></td>
<td>TINs are often created from grid or contour maps through the selection of sample points, as raw elevation data are not readily available. The advent of LIDAR data, however, may change this.</td>
<td>Elevation data are readily available in grid formats.</td>
</tr>
<tr>
<td><strong>Computational Efficiency</strong></td>
<td>Triangulation requires more computational power then creating DEMs; calculating slope and aspect are easier with TINs.</td>
<td>Can perform map algebra functions (computing values in one layer against values in another layer). This is particularly important when ancillary data, such as land use/land cover or soils data are needed.</td>
</tr>
<tr>
<td><strong>Storage Efficiency</strong></td>
<td>Point-for-point, requires more storage (every x,y,z maintained), particularly if topology is maintained; point data storage can be reduced as fewer points are required to create a TIN of the same or better accuracy as a DEM.</td>
<td>Storage structure straightforward (horizontal implicit with grid structure); storage is less efficient for homogeneous areas (however, algorithms are available for compressing data).</td>
</tr>
<tr>
<td><strong>Sample Points</strong></td>
<td>Directly incorporates sample points in triangulation (interpolation only used for queries); better terrain definition.</td>
<td>Generalizes terrain surface through interpolation to grid structure; some precision/definition lost.</td>
</tr>
<tr>
<td><strong>Interpolation</strong></td>
<td>Interpolation (natural neighbor) is used only to query unsampled points; interpolation is not used to create TIN.</td>
<td>Interpolation is used to create grid from irregular points; more robust interpolation approaches available than for TIN.</td>
</tr>
<tr>
<td><strong>Stream Extraction</strong></td>
<td>Trigonometric nature conducive to mapping flow (don’t need to worry about multiple flow direction or D8 vs. D-infinity); triangulation incorporates breaklines as triangle edges; difficult to determine if proper drainage exists over large regions; cannot automatically determine stream origination.</td>
<td>Model is more mature; more approaches available and tested; must cope with accumulation flow thresholds and multiple vs. single flow direction approaches; easier to determine if proper drainage exists.</td>
</tr>
</tbody>
</table>
### Table 5.1 Continued

<table>
<thead>
<tr>
<th>Criteria</th>
<th>TIN</th>
<th>DEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease/Convenience of use</td>
<td>Most data are in grid formats that must be converted to TIN; fewer models are available for use; cannot use map algebra functions.</td>
<td>More algorithms and models available; user knowledge needed to use interpolation techniques appropriately, must determine accumulation flow thresholds, multiple vs. single flow, etc.</td>
</tr>
</tbody>
</table>
6.0 References Cited


Environmental Modeling Research Laboratory (EMRL) of Brigham Young University, WMS v7.0 Tutorials, Provo, Utah, 2002.


Poiker, T. K., Unit 39, NCGIA Core curriculum in geographic information science. Simon Fraser University, 1990.


Automated Techniques Using GIS and LIDAR to Map Headwaters Stream Networks
In the Piedmont Ecoregion of North Carolina

1.0 Introduction

Headwaters stream networks serve as a critical hydrologic link between the surrounding landscape and larger, connecting stream outflows within a watershed. These networks drain extensive surface areas of land within the watershed that are not directly in contact with higher order stream channels. As changes in the adjacent land use and land cover occur from human and natural pressures, flows occurring in these lower order streams export diffuse pollution that has been aggregated from throughout the watershed (Meyer and Wallace, 2001; Peterson, et al., 2001).

Accurate stream maps are needed to investigate the characteristics, magnitude and mitigation of non-point source pollution at headwaters streams. Local, state, and federal agencies have relied on US Geological Survey 1:24,000 scale topographic maps for information on stream networks for planning, management, and regulatory programs related to streams. However, recent research in North Carolina has demonstrated the relatively low accuracy of US Geological Survey (USGS) 1:24,000 scale maps in depicting the presence and location of low order, headwaters streams (Darling, et al. 2002; Gregory, et al., In Review, 2004). In terms of stream length, the USGS topographic maps, on average, overestimate the presence of headwaters streams in the Coastal Plain by 31%, underestimate the presence of headwaters streams in the Piedmont by 25%, and in the Mountains, by up to 54% (Gregory, et al., In Review, 2004).

Currently, the very low accuracy of topographic maps in depicting first- and second-order streams places an inordinate burden on both regulatory agencies and the regulated
community to conduct joint field checks of stream locations. Improved headwaters streams maps would be useful to a wide variety of state agencies in planning activities, such as, implementing road construction projects and performing industrial recruitment efforts. Local governments would benefit from the maps in their regional planning efforts and application of riparian buffer regulations and ordinances. Thus, a new automated mapping approach for identifying headwaters stream networks would contribute substantially to state and regional planning and regulatory communities, as well as local regulated communities.

In this paper, I present a survey method developed for validating high-accuracy LIDAR data, compare interpolation methods used to generate topographic maps from these data; evaluate the effect of scaling on the generation of topographic and headwaters stream network maps; and assess the impact of using hydro-enforcement on modeling stream networks.

2.0 Background

Automated stream mapping approaches require accurate topographic maps, as gravity is the most prominent factor in the drainage of water from the landscape (Maune et al., 2001; Theobald and Goodchild, 1990). The advent of Light Detection and Ranging (LIDAR) technology has presented new opportunities in producing high-density, high-accuracy elevation data as a source for accurate topographic maps (Renslow, 2001; Maune et al., 2001; and Band, 1993). This increase in accurate data has enhanced the quality, definition and resolution of available topographic maps (Osborn, et al., 2001, Fowler, 2001). LIDAR also results in more accurate elevation maps in vegetated areas, because LIDAR needs only a single pulse to reach the ground whereas photogrammetry requires different lines of sight to
see the same points on the ground (NC Floodplain Mapping Program, 2003; Thompson and Maune, 2001; Renslow, 2001). The disadvantages of using LIDAR are: (1) LIDAR results in artifacts—erroneous bare-earth elevation determinations caused by returns from features such as, bridges, culverts and dams; (2) LIDAR pulses are fully absorbed by water, rendering water returns unreliable; (3) densely vegetated areas (e.g., forest, grass, shrub) can produce false bare-earth elevation data; and (4) post-processing removes questionable returns resulting in irregularly spaced points and sometimes large areas with no data (Renslow, 2001; Spinney, 2001; NC Floodplain Mapping Program, 2003; Thompson and Maune, 2001).

In a review of the use of LIDAR for stream mapping, I found that:

1. LIDAR data present new challenges in automatically extracting stream networks. While the generally higher density of the LIDAR bare-earth elevation points vastly improves interpolation results, voids in the terrain data caused by factors such as absorption and/or post-processing present interpolation challenges. LIDAR data also produce artifacts that can affect the flow of water as predicted by GIS stream network algorithms, requiring hydro-enforcement (stream centerlines used to enforce drainage, commonly referred to as “breaklines”) (Renslow, 2001; Spinney, 2001; NC Floodplain Mapping Program, 2003; Thompson and Maune, 2001).

2. TINs produce more precise topographic maps. TINs maintain the LIDAR elevation points by incorporating them into the triangle vertices. Conversely, DEMs interpolate the mass elevation points, resulting in a loss of precision. Thus, TINs better represent the natural variability of the terrain and retain linear structures (e.g., stream channel lines or breaklines) better then DEMs (Kumler, 1990; Nelson, et al., 1999; Dodson and Li, 2000; Maune et al., 2001; Wang, 2001).
3. DEMs are better for automatically extracting headwaters streams. Both the DEM and TIN will produce artifacts obstructing drainage (e.g., pits, flat areas, false dams). Because of its irregular structure, it is more difficult to automatically correct drainage problems in a TIN than in a DEM (Olivera, 2002, EMRL, 2002). In addition, TINs do not offer an automated way to determine stream origination (EMRL, 2002).

4. The interpolation method and the resolution used to generate the topographic map can impact the accuracy of the map. Different interpolation methods produce different results, particularly in areas where there are sparse data. Some of the more complex interpolation methods (e.g., kriging) require more knowledge by the user and are computationally demanding, but are expected to perform better in modeling terrain.

Related to this issue is the question of scale: at what resolution (i.e., cell size) is the headwaters stream channel definition lost? The resolution drives the storage and computational demands of the interpolation method and can impact the selection of which interpolation method can be used (Band, 1993; Nelson, 1999; Maune et al., 2001). The resolution also impacts the amount of processing time required for the stream extraction programs.

5. Hydro-enforcement of the DEM enhances the accuracy of the stream network map. Several authors noted that hydro-enforcement of the TIN or DEM (e.g., incorporating the stream centerlines into the TIN triangle edges or overlaying them onto the DEM) contribute significantly to improving stream extraction results (Theopald and Goodchild, 1990; DeVantier, B. and A. Feldman, 1993; Maune et al., 2001; Thompson and Maune, 2001; NC Floodplain Mapping Program, 2003). Breakline data, however, are expensive to develop, as they must be derived from orthophotographs.
6. The physical processes relevant at headwaters stream scales are different than watershed scales typically modeled. Physical characteristics, such as overland flow, that are “averaged-out” at coarse scales (e.g., basin or watershed) become critical at fine scales (e.g., headwaters stream catchments). Thus, accurately modeling hydrologic characteristics such as stream origin and hillslope flow may require more sophisticated methods (e.g., diffuse flow) than the standard algorithms commonly used (Tarboton and Ames, 2001; Tarboton, 1997).

3.0 Objectives

The conclusions from the literature search in Part I of this thesis formed the basis for the research questions addressed by this study:

- What is the accuracy of the LIDAR surface elevation points?
- When using densely spaced LIDAR data, does the interpolation method used to create a DEM make a difference in the accuracy of the DEM?
- At what resolution do you begin to lose channel definition of headwaters streams, thereby affecting the production of headwaters stream maps?
- Does using breakline data to hydro-enforce TINs and DEMs make a difference in the production of headwaters stream maps?

Four objectives were developed from these research questions and are the basis for the organization of this thesis: (1) evaluate the accuracy of LIDAR surface elevation points; (2) compare the accuracy of topographic maps generated using various interpolation methods; (3) evaluate the effect of scaling on the topographic and stream map accuracy; and
(4) assess the impact of using hydro-enforcement to extract stream networks (Figures 3.1 - 3.3).

4.0 Methods

Study site. The forested site selected to examine the study objectives is located in Falls Lake State Park, Wake County, North Carolina (Figures 4.1 and 4.2). The study results and conclusions are limited in applicability to the selected study site, and the predominantly crystalline felsic, fluvial eroded terrain of the Piedmont Ecoregion. The study site was selected based on; (1) the availability of LIDAR data, (2) the accessibility of the site for laying survey benchmarks and collecting field data, (3) the relatively natural, undisturbed conditions of the watershed, and (4) the variation of catchment sizes and profiles within the watershed (e.g., steepness of gradients, number of intermittent and perennial streams).

The Piedmont is composed of mostly gneiss and schist rock intruded by granitic plutons, with a thin layer of saprolite (Griffith, et al., 2002). The most prevalent soil types mapped in the study area are Madison sandy loam (fine, kaolinitic, thermic Typic Canhapludults), Cecil (gravelly sandy loam thermic Typic Canhapludults), and Wilkes (loamy, mixed, active, thermic shallow Typic Hapludults) (Cawthorn, 1970). The soil in a small catchment selected for intensive study is mapped as Madison sandy loam. Most of the soils are noted for their high occurrence of pebbles and cobbles, high susceptibility to erosion, and rapid runoff rates. Notwithstanding the rapid runoff rates, infiltration is generally rated as "good". Many of the soils are not suitable for cultivating--a fact discovered during the collapse of the agriculture economy in this area nearly 70 years ago.
Forest vegetation of the study area consists of natural forests that resulted from old field succession after agricultural abandonment approximately 70-90 years ago. Overstories in different stands vary from mixed pine to mixed pine-hardwood to mixed hardwood. Steep slopes adjacent to the headwaters streams generally have mixed hardwood stands that have much larger and older trees than the mid and upper slopes. The midstories of all stands are mixed hardwoods. The forest vegetation of the small, intensively studied catchment is representative of the area. The mid and upper slopes have two stand types, mixed pine and mixed hardwood-pine. In the mixed pine stands, the dominant overstory species are loblolly pine (Pinus tadea L.) and shortleaf pine (Pinus echinata Mill.), with a few scattered Virginia pine (Pinus virginiana Mill.). The largest trees (loblolly pine) are 16-18 inches diameter-at-breast-height (dbh). In the mixed hardwood-pine stands, the dominant overstory species are red maple (Acer rubrum L.), yellow poplar (Liriodendron tulipifera L.), sweetgum (Liquidambar styraciflua L.) and loblolly pine with a few white oaks (Quercus alba L.) and shortleaf pines. The largest trees (yellow poplar) are 18-20 inches dbh. The midstories of both stand types are dominated by red maple, sweet gum, and yellow poplar with a few white oaks. The overstory of mixed hardwood trees on the steep lower slopes are dominated by white oak, yellow poplar, sweet gum, and red maple with a few large black oaks (Quercus velutina Lam.). The largest trees are 28-30 inches dbh.

**Modeling approach.** I used the Environmental Systems Research Institute's (ESRI’s) ArcGIS version 8.1 (2002) for all modeling conducted in the study. All models were applied to the study area to produce topographic and headwaters stream maps, and these various maps and techniques were evaluated using field collected survey and GPS data. I used a larger study area (approximately 239 acres) to assess the accuracy of the LIDAR
surface elevation points and the modeled topography (Figure 3.2) and a smaller catchment area (approximately 26 acres) to assess the stream mapping results (Figure 3.1).

In order to take advantage of the strengths of both the TIN and DEM topographic modeling approaches, three topographic mapping methods were used for this study: (1) TIN-only for model validation; (2) combined TIN/DEM approach for assessing the superior hydro-enforcement techniques available with a TIN, while still incorporating the automatic stream extraction methods associated with the DEM; and (3) DEM-only approaches to assess various interpolation methods, the impact of scale on channel definition, and the benefits of using hydro-enforcement to extract stream networks. The term “stream extraction” is used throughout this paper to indicate the modeling of flow paths, rather than the modeling of channel morphology or other hydraulic processes that could be assumed with the term “stream mapping”.

**Data.** LIDAR surface elevation points and breakline data used for this study were obtained from the North Carolina Cooperating Technical State Floodplain Mapping Program (NC State Floodplain Mapping Program) (2003). Ancillary data used in the comparisons of the breakline data and the extracted stream networks included the medium and high resolution National Hydrography Dataset (NHD) available for the upper Neuse River (1999), Wake County Hydrography Lines (2000), the USGS 1:24,000 Topographic Digital Raster Graphs (DRGs) (1994), and 1999 digital orthophotography (Wake County North Carolina, 2000).

The medium resolution NHD integrates the content of the USGS 1:100,000 Digital Line Graph (DLG) hydrography data and the reach-related information for the US Environmental Protection Agency Reach File Version 3.0 (RF3). The high resolution NHD
uses the USGS 1:24,000 topographic map, and is supplemented by additional data (e.g., Digital Orthophoto Quadrangles (DOQs)) where available. The medium resolution datasets are available for the State of North Carolina, and the high resolution datasets are available for a limited number of counties in the State, including Wake County, the location of the study site (NHD, 1999).

The Wake County North Carolina hydrography lines (2000) are derived from 1:12,000 aerial photography. The Wake County color digital orthophotography (1” = 200’ for rural; 1” = 100’ for urban) are from a 1999 county-wide aerial flight (Wake County North Carolina, 2000).

4.1 Objective 1: Evaluation of LIDAR Surface Elevation Points

The accuracy of the LIDAR-derived surface elevation points was published by the NC Floodplain Mapping Program to be within .66 ft. (20cm) for coastal areas and .82 ft. (25cm) for inland areas (2002a). Because of the high accuracy of the LIDAR data, I used survey-grade techniques for collecting the ground-truth data. I established control benchmarks along the ridge of the study catchment using rigorous surveying techniques, and then surveyed the headwater stream catchment using the newly-established control benchmarks as the starting point. In addition, I collected horizontal coordinates along the study stream channel for assessing the modeled stream network.

Establishment of control benchmarks. I paid particular attention to establishing the nine control benchmarks located along the ridgeline of the study area (Figure 4.1). I used three survey techniques to achieve the highest horizontal and vertical accuracy that I could for these benchmarks; centimeter accuracy differential GPS, total station survey/traverse, and
differential leveling. Survey-grade GPS can produce vertical measurements to within +/- 0.07 ft. (2cm), provided that National Geodetic Survey (NGS) guidelines are strictly followed when taking these measurements (NGS, 2002). In compliance with these guidelines, two monuments, Ridge and Emorys, were selected as base stations. The Ridge and Emorys monuments were permanently set by the NGS using second-order horizontal standards (relation of specific accuracy to the coordinates of all other points in the horizontal network is at least 1:20,000) and second-order, class 0 vertical accuracy standards for the Ridge monument (accurate to within +/- 0.0028 ft. [8.4 mm]) (NGS, 1985). Two Leica Ski-Pro v.2.1 GPS units were used, with one unit stationed on the NGS monument, and a second unit placed at the site of the control benchmark. This process was repeated twice (on separate days, during different times of the day), for each of the control benchmarks established in this manner, and resulted in vertical measures accurate to within +/- .33 ft. (10 cm). I was only able to set four of the nine control benchmarks using this technique because of the demanding satellite-reception thresholds required by the survey-grade GPS units.

Thus, I turned to the Total Station Survey/Traverse (TSS/Traverse) technique to establish the horizontal locations of the remaining control benchmarks. A traverse is a series of consecutive surveyed lines whose lengths and directions have been determined from field measurements (Brinker and Wolf, 1977). With the help of a certified surveyor, I used a “closed traverse” method consisting of line-of-site horizontal angle and distance measurements (forward and backward) at each control benchmark along the ridgeline, and then a repeat of these measurements back to the beginning point. This method provided redundant checks on the measured angles and distances. Horizontal measures (vertical measures were not taken using this method) were adjusted using least squares calculations.
and resulted in a standard error of +/- 0.033 ft. (1 cm).

Differential Leveling was used to measure vertical distance at each of the control benchmarks. Differential Leveling is a trigonometric-based method used to derive elevation between two points by measuring the inclined distance and the vertical angle to one point from a horizontal plane through the other (Brinker and Wolf, 1977). I used a permanent monument (FN-49 with a vertical measure accurate to within +/- 0.004 ft. [12 mm]) established by the U.S. Army Corps of Engineers (USA COE, 2002) as the starting point of our leveling. Using line-of-site instrumentation, I (with help from a certified surveyor) measured the inclined distance and vertical angle (forward and backward between points) from the USA COE monument to the last benchmark at the study site. I then repeated the measurements back to obtain a vertical accuracy of +/- 0.07 ft. (2 cm) for the control benchmarks.

Survey of catchment and measurement of stream points. The catchment located at benchmark z was chosen for detailed study because of the narrow, yet defined, stream channel and the variable slope of the catchment (gradual on the northern slope of the channel, steep on the southern slope) (Figure 4.2). In order to assess how well the models captured this terrain variability, I surveyed a transect line across the catchment, beginning at the ridgeline, extending down the catchment to the main headwater stream channel, and then up the other side of the catchment. The survey plan developed for obtaining the transect line is unique in that it maximizes accuracy, yet optimizes the number of points taken across the transect. To maximize accuracy, I set benchmarks using the TSS/Traverse technique in a loop down through the catchment, beginning and ending at control benchmark z. I then took several measurements along the transect from the appropriate catchment benchmarks and
calculated the standard error for the transect vertical measurements as +/- .26 ft. (8 cm).

The origin of the first order intermittent stream in the study catchment was
determined in the field using the NC Division of Water Quality Stream Classification
methodology (Darling, et al., 2002). Differential mapping-grade GPS was used to measure
the horizontal coordinates of the stream origin and the major points of directional change
along the main channel (no vertical measures were taken with the mapping-grade GPS)
(Figure 4.2). Horizontal accuracy of mapping-grade GPS is from one to two meters (Jenkins,
2002).

**Comparison of LIDAR TIN elevation data to survey points.** Because survey
points rarely overlay exactly on LIDAR sample points, I generated a TIN from the LIDAR
data (without the incorporation of breaklines) and used this TIN for the comparison (Daniel
and Tennant, 2001). Comparisons between the LIDAR data (represented by the TIN) and
survey elevation measurements were made by overlaying the survey points at the same x,y
coordinates on the TIN. The survey points were magnified (using the ArcMap “zoom”
function), and the elevation values were recorded for each of the layers of interest in an Excel
spreadsheet. I used Excel to assess how well the interpolated points aligned with the survey
points using graphs, and to calculate the root mean square error (RMSE) for each of the layer
pairs. RMSE was calculated using the following formula:  \[ \text{RMSE}_z = \sqrt{\frac{\sum (z_{\text{data } l} - z_{\text{check } l})^2}{n}} \], where "z_{\text{data } l}" is the vertical coordinate of the \( l \)th check point in the dataset, "z_{\text{check } l}" is
the vertical coordinate of the \( l \)th check point in the independent source of higher accuracy, "n"
is the number of points being checked, and "l" is an integer from 1 to n.
4.2 Objective 2: Comparison of Topographic Maps Generated using Various Interpolation Methods

I used LIDAR data to produce accurate topographic maps by generating DEMs using four different interpolation techniques and comparing the DEMs to survey points. The four interpolation methods, inverse distance weighted (IDW), spline, kriging and natural neighbor, are the most widely used methods for generating DEMs and TINs.

**Interpolation methods.** IDW (fixed and variable), spline (regularized and tension), and kriging (layer-to-grid and points-to-grid) interpolation methods were applied directly to the original LIDAR data to generate 20 ft. resolution DEMs. In addition, the TIN constructed from LIDAR data was converted to a grid using the only available interpolation method supported by ESRI’s ArcMap Spatial Analyst extension--natural neighbor.

The interpolated grid elevation values were then compared to the field-collected survey elevation data. A grid resolution of 20 ft. (cell size of 20x20 sq. ft.) was selected for this comparison because the LIDAR data density produced by the NC Floodplain Mapping Program averaged at least one point for every 20x20 sq. ft. (2003). A total of seven DEMs were created, using the interpolation methods and settings depicted in Table 4.2.1.

<table>
<thead>
<tr>
<th>Data</th>
<th>Method</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIN (w/o breakline data) converted to DEM</td>
<td>Natural Neighbor</td>
<td>Z factor = 1</td>
</tr>
<tr>
<td>LIDAR data</td>
<td>Inverse Distance Weighted</td>
<td>Radius = fixed, power 2, distance = 200’</td>
</tr>
<tr>
<td>LIDAR data</td>
<td>Inverse Distance Weighted</td>
<td>Radius = variable, power = 2 # of points = 12</td>
</tr>
</tbody>
</table>
### Table 4.2.1 Continued

<table>
<thead>
<tr>
<th>Data</th>
<th>Method</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIDAR data</td>
<td>Spline – Regularized</td>
<td>Weight = .1, # of points = 12</td>
</tr>
<tr>
<td>LIDAR data</td>
<td>Spline – Tension</td>
<td>Weight = .1, # of points = 12</td>
</tr>
<tr>
<td>LIDAR data</td>
<td>Kriging – continuous kriged layer-to-grid</td>
<td>Ordinary, Spherical Model, Lag = 100, # lags = 20</td>
</tr>
<tr>
<td></td>
<td>(Geostatistical Analyst)</td>
<td></td>
</tr>
<tr>
<td>LIDAR data</td>
<td>Kriging – points-to-grid (Spatial Analyst)</td>
<td>Ordinary, Spherical Model, Lag = 100, # lags = 20</td>
</tr>
</tbody>
</table>

The Natural Neighbor interpolation method available in Spatial Analyst (ESRI, 2002) was used to convert the irregularly spaced TIN to a regularly spaced grid structure. A Thiessen polygon (a boundary created from assigning locations to the nearest vertex) (Poiker, 1990), is determined for the unsampled grid point, and overlain over the Thiessen polygons resulting from the sampled points. The area of the overlap is used to weight the influence of the local sampled points, resulting in a surface with smoothly changing gradients (Watson, 1992, Maune et al., 2001).

I used the interpolation options available through ESRI’s Spatial Analyst (IDW, spline and kriging) to interpolate directly from the irregularly spaced LIDAR points to a regularly spaced grid structure,. With the IDW function, a fixed radius (circle of a fixed distance to find input points) or variable radius (adjusts the radius for including input points to a minimum number of points) can be selected. For both methods, a higher power puts more emphasis on the nearest points, creating a more detailed, but less smooth surface. A fixed radius is recommended when the source data points are plentiful and uniformly spaced, whereas, a variable radius is recommended when the points are sparsely and/or randomly spaced.
placed (ESRI, 2002). Because the source input data are dense, but randomly spaced, both methods were compared.

Two spline methods are available through Spatial Analyst, regularized and tension. Regularized spline creates a smooth surface and slope, while tension tunes the stiffness of the surface based on surrounding points—similar to a rubber sheeting effect. The weight option affects the regularized and tension methods differently. The higher the weight for regularized spline, the smoother the surface, whereas, the higher the weight for tension spline, the coarser the surface and the more the range of data values come into play (ESRI, 2002). While regularized spline is more appropriate for modeling terrain where channel definition in hillslope areas is important, both the regularized and tension spline were compared in this study.

Kriging offers two main approaches: Ordinary and Universal. Ordinary kriging assumes the constant mean is unknown and must be calculated from the data, whereas, universal kriging assumes that there is an overriding trend in the data that can be modeled by a deterministic function (polynomial). Similar to IDW, kriging weights the surrounding sample points to derive a prediction for each location based on the distance between the measured points and the predicted location. In kriging, the weights are not only based on the distance between points, but also on the variation between measured points as a function of distance. This is referred to as spatial variation and is analyzed using variograms, which plot the variance of paired sample measurements as a function of distance between samples. A parametric model (e.g., exponential, spherical, gaussian) is then selected based on which model best fits the variogram. This process is similar to regression analysis in that a continuous curve is being fitted to the data points in the variogram. After a suitable
variogram model has been selected, kriging creates a continuous surface for the entire study area using weights calculated based on the variogram model and the values and location of the measured points (US EPA, 2003).

Two kriging methods available through ESRI’s ArcMap (2002) application were compared; LIDAR surface elevation points were interpolated to a grid structure through the kriging option in the Spatial Analyst extension, and a continuous layer was generated through the kriging option in the Geostatistical Analyst extension and then exported to a grid structure (ESRI, 2002).

**Comparison of DEMs and survey points.** I compared the LIDAR data (represented by the TIN), DEMs and survey points by overlaying the survey points at the same x,y coordinates on the TIN and DEMs. I magnified the survey points (using the ArcMap “zoom” function), and the elevation values were recorded for each of the layers of interest in an Excel spreadsheet. Excel was used to assess how well the interpolated points aligned with the survey points using graphs, and to calculate the root mean square error (RMSE) for each of the layer pairs (see Section 4.1 for the RMSE formula).

**4.3 Objective 3: Effect of Scaling on the Accuracy of Topographic Maps and Extracted Stream Networks**

I examined the impact of scaling to finer and coarser resolutions on the modeled terrain definition and the extraction of stream networks. The interpolation methods (regularized spline and natural neighbor) that produced the best results in the previous step were used to examine the effect of scaling. IDW was also evaluated as I was interested in assessing this more simplistic method. I used these interpolation methods to generate 10 ft.,
20 ft., 60 ft. and 90 ft. resolution DEMs for a total of 12 topographic maps (three interpolation methods x four resolutions). These twelve maps were used to extract twelve stream networks at 10 ft., 20 ft., 60 ft. and 90 ft. resolution to determine the best combination of the interpolation method and resolution.

**Comparison of DEMs generated at different resolutions.** I compared the twelve topographic maps to the surveyed elevation points to assess the terrain definition lost or gained at the varying resolutions. I compared the LIDAR data (represented by the TIN), DEMs and survey points by overlaying the survey points at the same x,y coordinates on the TIN and DEMs. The survey points were magnified (using the ArcMap “zoom” function), and the elevation values were recorded for each of the layers of interest in an Excel spreadsheet. Excel was used to assess how well the interpolated points aligned with the survey points using graphs, and to calculate the root mean square error (RMSE) for each of the layer pairs (see Section 4.1 for the RMSE formula).

**Comparison of stream networks extracted at different resolutions.** ESRI’s ArcHydro extension was used to extract the stream networks from the twelve DEMs generated at the various resolutions for a total of twelve stream network maps. An average catchment size of 80,000 sq. ft. was used to calculate the accumulation threshold for each of the stream networks. The various stream networks were compared and the differences were noted. In addition, the field-collected GPS points and the Wake County hydrography lines (2000) were overlain on the stream network maps to assess the quality lost or gained through scaling.
4.4 Objective 4: Assessment of Hydro-Enforcement on Extracting Stream Networks

I examined; (a) the accuracy of the source breakline data, (b) the impact of using breakline data to extract headwaters streams, and (c) the alignment of the extracted stream networks with currently available stream maps.

**Assessment of Breakline Data.** The breakline data used for this study were derived from photogrammetry flown concurrently with the LIDAR data collection, or from existing orthophotography (NC Floodplain Mapping Program, 2002b). While metadata were available for the breakline data, the accuracy information provided focused on the assignment of LIDAR elevation values to the digitized 2-D vector lines. No information was provided on the positional accuracy of the digitized vector lines. Because errors in the source breakline data can significantly impact the mapping of streams, the accuracy of the source breakline data were checked against the Wake County 1999 digital orthophotography (2000). The breakline data were also compared to high (1:24,000) resolution NHD (1999), the Wake County hydrography lines (2000) and the USGS 1:24,000 Topographic DRG (1994).

**Comparison of stream networks.** I used four DEMs (20 ft. resolution) as base maps for extracting the stream networks: two DEMs were generated from the TINs (created with and without breakline data using the natural neighbor interpolation method) and two DEMs were generated directly from the LIDAR data using the IDW and regularized spline interpolation methods. The kriging method was not used for this comparison because of the relatively poor results achieved in the comparison of the interpolation methods and the complexity of the method.

ESRI's ArcHydro extension was used to extract the stream networks. ArcHydro uses the same basic algorithms used in other stream extraction methods, but incorporates an
improved DEM surface reconditioning algorithm called AGREE (Hellwegner, 1997). In addition to filling spurious pits, the AGREE algorithm incorporates breaklines into the DEM and adjusts the elevation of an area around the stream to allow a gradual descent to the channel.

I applied ArcHydro to each of the four DEMs to incorporate breaklines (using the AGREE algorithm), fill spurious pits, determine flow direction, calculate flow accumulation, and extract the drainage network. (It should be noted that the stream centerlines had to be separated from the waterbody outlines in the breakline file before this data could be used for hydro-enforcement.) ArcHydro was also applied to each of the four DEMs without incorporating breaklines. A total of eight stream maps (four with hydro-enforcement and four without hydro-enforcement) were produced for this comparison.

Table 4.4.1 outlines the interpolation method and hydro-enforcement approach used to extract the various headwater stream drainages and provides a key to reading the figures in this section.

<table>
<thead>
<tr>
<th></th>
<th>TIN-to-grid w/breaklines</th>
<th>TIN-to-grid w/o breaklines</th>
<th>Grid generated using IDW</th>
<th>Grid generated using Regularized Spline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ArcHydro (w/ AGREE)</strong></td>
<td>TINgrid with breaks/with AGREE</td>
<td>TIN-to-grid no breaks/with AGREE</td>
<td>IDW with AGREE</td>
<td>Spline- reg with AGREE</td>
</tr>
<tr>
<td><strong>ArcHydro (w/o AGREE)</strong></td>
<td>TINgrid with breaks/no AGREE</td>
<td>TINgrid no breaks/no AGREE</td>
<td>IDW no AGREE</td>
<td>Spline- reg no AGREE</td>
</tr>
</tbody>
</table>

The eight stream maps were compared to note differences in the modeled lower-order
(headwaters) and higher-order stream networks. In addition, the networks were compared to the stream GPS points, the medium (1:100,000) and high (1:24,000) resolution USGS NHD (1999), the Wake County hydrography lines (2000), Wake County 1999 orthophotography (2000), and the USGS 1:24,000 Topographic DRG to assess the quality lost or gained through the modeling techniques applied in this study.

5.0 Results and Discussions

5.1 Objective 1: Evaluation of LIDAR Surface Elevation Points

Comparison of LIDAR elevation data. The total RMSE for the comparison of the LIDAR elevation data (as predicted by the TIN) to the survey elevation data was 1.2 ft (37 cm). Examining the location of the discrepancies along the elevation transect reveals that most of the differences between the LIDAR and survey data were in the channel area (Figure 5.1.1). The LIDAR TIN failed to show the elevation of the channel bottom, but does an excellent job at depicting the lowest area of the valley where the stream is located.

The accuracy of the surface elevations as determined from the LIDAR TIN was comparable to the published elevation accuracy. The published accuracy of the LIDAR data is qualified as .82 ft. (25 cm) with “5% of points having the largest error removed” (NC Floodplain Mapping Program, 2002a). Daniel and Tennant (2001) call this approach the “95 percentile” method and explain that 5% of the outliers, as qualified by the 3-sigma rule (standard deviation x 3), are removed before calculating RMSE to obtain a normal distribution. To remain consistent with this approach, the 95 percentile method was applied to the RMSE calculation for the study. The recalculated RMSE was .78 ft. (24 cm) and is within the published error of .82 ft. (25 cm).
This is particularly noteworthy as the survey points collected for the NC State Floodplain Mapping Program were obtained on flat or uniformly-sloping terrain for 16.4 ft. (5 m) in all directions (2002a). I did not consider consistent elevation in the selection of survey points, and purposely collected a transect across variable terrain. In addition, the North Carolina Floodplain Mapping Program collected a minimum of 20 points in each of five landcover types. The published RMSE for the LIDAR data was calculated on the total variance for all landcover types (NC Floodplain Mapping Program, 2002a). This study was conducted in one landcover type (forested).

5.2 Objective 2: Comparison of Topographic Maps Generated Using Various Interpolation Methods.

The elevation values produced by the regularized spline DEM were closest to the surveyed elevations with a RMSE value of 1.4 ft. (Table 5.2.1). The elevation values produced by the natural neighbor (TIN-to-grid) DEM also aligned closely with the surveyed elevations with a RMSE value of 1.6 ft. The RMSE for the elevation differences rises sharply to 1.9 ft. with the kriged DEM, and the IDW DEM performed the weakest with a RMSE of 2.1 ft.

<table>
<thead>
<tr>
<th>Interpolation Method</th>
<th>RMSE (no points removed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIDAR TIN Surface Elevation Points</td>
<td>1.2 ft.</td>
</tr>
<tr>
<td>Regularized Spline DEM (20 ft.)</td>
<td>1.4 ft.</td>
</tr>
<tr>
<td>Natural Neighbor DEM (20 ft.)</td>
<td>1.6 ft.</td>
</tr>
</tbody>
</table>
Table 5.2.1 Continued

<table>
<thead>
<tr>
<th>Interpolation Method</th>
<th>RMSE (no points removed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kriged DEM (20 ft.)</td>
<td>1.9 ft.</td>
</tr>
<tr>
<td>Inverse Distance Weighted DEM (20 ft.)</td>
<td>2.1 ft.</td>
</tr>
</tbody>
</table>

**IDW.** Both the fixed and variable radius options were compared for the IDW interpolation method. For the variable option, a minimum of 12 and 20 points were used, with no difference in the results. The RMSE was the same for the fixed and variable methods, but the fixed radius option (set at 200 ft.) did not interpolate across the waterbody, leaving zero values for these areas. Not only is the estimated elevation for the waterbody not calculated, but the result creates problems when performing stream extraction. The RMSE for both methods was 2.2 ft.

**Spline.** The RMSE for the regularized and tension spline was one of the best at 1.4 ft. While there was no difference between the regularized and tension spline methods in the RMSE calculation or the alignment of the modeled points with the field-collected data, the tension spline method calculated negative values for the lake areas. These negative values created problems during the stream extraction process (see discussion in Section 5.4).

**Kriging.** Kriging was the most computationally demanding interpolation method used for this study. The generation of the layer using kriging required substantially more computing power and time (approximately 20 minutes as compared to under 5 minutes for the other techniques) and required that the data be reduced in order to successfully run the models. In addition, knowledge on fitting the parametric model to the variogram was essential in producing good results. I consulted with a senior statistician within EPA
(personal comments, David Holland, Office of Research and Development, March, 2004) to correctly apply the model parameters.

The continuous layer generated by Geostatistical Analyst produced the best results as compared to the survey data (RMSE = 1.4 ft.). Because a continuous layer cannot be used for stream extraction, the layer was exported to a DEM. Converting the continuous layer to the grid structure resulted in a loss of accuracy (RMSE of the converted layer-to-grid = 2.3 ft.). The LIDAR elevation points were also directly interpolated to a grid structure using the kriging option in Spatial Analyst, resulting in a RMSE of 1.9 ft.

5.3 Objective 3: Effect of Scaling on the Accuracy of Topographic Maps and Extracted Stream Networks

Comparison of DEMs generated at different resolutions. The LIDAR surface elevation points were used to generate DEMs at 10 ft., 20 ft., 60 ft. and 90 ft. resolution grids using the regularized spline, natural neighbor and IDW interpolation methods. I anticipated that a smoother surface would result at the finer 10 ft. resolution, particularly with the smoothing functions inherent in the regularized spline method. I found, though, that there was little difference between the 10 ft. and 20 ft. resolutions, regardless of the interpolation method used (Figure 5.3.1). A large difference was seen at the 60 ft. resolution; the 2 ft. headwater stream channel became a 60 ft. depression, demonstrating a critical loss of terrain definition at this resolution. At the 90 ft. resolution, the definition of the entire headwater catchment is lost (Figure 5.3.2).

Comparison of stream networks extracted at different resolutions. Stream networks were extracted from the 10 ft., 20 ft., 60 ft. and 90 ft. resolution grids using ESRI’s
ArcHydro extension. I found very little difference between the extracted stream networks using a 10 ft. and 20 ft. resolution DEMs (Figure 5.3.3). At the 60 ft. resolution, however, the stream drainage lines become more linear (grid-bias is more evident) and have more occurrences of drainage interruptions (Figure 5.3.4). These same problems were more pronounced in the stream network extracted from the 90 ft. resolution DEMs.

5.4 Objective 4: Assessment of Hydro-Enforcement on Extracting Stream Networks

I examined the breakline data used for hydro-enforcement and the impact of using hydro-enforcement on the resulting extracted stream network. The study was designed to use the interpolation methods with the best RMSE resulting from the comparison (natural neighbor and regularized spline). However, the smoothing function inherent in the regularized spline interpolation technique, smoothed the elevation across the Falls Lake fingers to negative values. This created poorer results in the stream extraction for the lake areas. Thus, I also extracted stream networks from DEMs generated using the IDW interpolation method.

Assessment of breakline data. The breakline data were compared to medium (1:100,000) and high (1:24,000) resolution NHD (1999) and Wake County Hydrography Lines (2000), and overlain on recent Wake County orthophotography (North Carolina, 2000). The stream centerlines from the breakline data aligned extremely well with the streams imaged in the orthophotography (Figure 5.3.5). Likewise, the medium (1:100,000) and high (1:24,000) resolution NHD stream centerlines aligned well with the streams shown on the USGS 1:24,000 Topographic Map, the known source for these data. It is important to note, however, that the location of the streams on the NHD map is different than on the
orthophotography (Figure 5.3.6). This finding is critical as the use of the NHD to hydro-enforce the DEM forces the stream network to be misaligned when compared to the orthophotography.

**Comparison of stream networks.** The use of the DEM generated from the hydro-enforced TIN (Figure 5.3.7) enhanced stream extraction in one small way; artifacts along the lake centerlines are apparent in the drainages generated without the hydro-enforced TIN and are non-existent in the drainage generated with the hydro-enforced TIN (Figure 5.3.8). This is the case only when the AGREE algorithm is also used to overlay the stream centerlines on the DEM. There was no advantage of using the DEM generated from the hydro-enforced TIN when the AGREE algorithm was *not* applied.

The model extracted the headwaters stream networks with only minor differences regardless of whether the DEM was hydro-enforced or not. The most significant differences between the various drainages were in the higher-order streams or breakline area. The use of the AGREE algorithm to hydro-enforce the DEM produced more straightforward drainages in the lake area than the drainages extracted without overlaying the stream centerlines (Figure 5.3.9). Without hydro-enforcement, the regularized spline grid resulted in particularly confused drainages in the finger lake area. This is likely due to the negative elevation values in the lake areas.

While the extraction of the headwaters stream network was substantially the same, regardless of the technique, the hydro-enforced TIN processed with the AGREE algorithm (Table 4.4.1: “TINgrid with breaks/with AGREE”) produced the best results in the higher order stream network. The “TINgrid with breaks/with AGREE” stream network was compared to the Wake County hydrography lines to validate the modeled stream network.
results (Figure 5.3.10). Both the stream maps were compared to the Wake County orthophotography and the field-collected GPS stream points (Figure 5.3.11). The comparison revealed that the modeled drainage does a better job than the photo-derived Wake County hydrography in mapping the actual stream in the study catchment.

Finally, the “TINgrid with breaks/with AGREE” drainage was compared to the USGS 1:24,000 Topographic DRG to visually assess the quality lost or gained with this technique. The modeled drainage does not align with the topographic drainage contour lines in several places. This finding is consistent with the assessment of the breakline data in Section 5.4 that showed that the NHD drainages do conform to the 1:24,000 Topographic DRG, but do not align to the 1999 orthophotography, Wake County hydrography lines, or the NC State Floodplain Mapping Program breakline data. For the areas measured, the difference between the topographic drainage contour lines and the modeled drainage was, on average, off by 110 ft. (Figure 5.3.12).

Figure 5.3.13 shows the total stream length for each of the compared methods using the AGREE algorithm. The total stream length is nearly the same (approximately 230,000 ft.) for the stream networks produced by Wake County, and the methods used in this study (except that the regularized spline method is slightly higher due to the confused drainages in the lake area). It is critical to note the additional stream length generated by the stream extraction process beyond the currently available NHD (132,608 ft. for high resolution and 25,824 ft. for medium resolution) and NC State Floodplain Mapping stream centerline data (123,072 ft.). This is particularly relevant when considering that the NHD – medium resolution stream centerline data are the only breakline data currently available for the entire State.
6.0 Summary and Conclusions

For the terrain associated with this study site in the Piedmont geographic region, and the dense LIDAR coverage available for this site, LIDAR produced more accurate elevation maps (elevation accuracy within 1.6’) than the currently available USGS 7.5 minute DEMs (elevation accuracy within 49’). While the spline interpolation method performed the best overall with a RMSE of 1.4 ft., the nearest neighbor technique (RMSE = 1.6 ft.) produced the best topographic map for extracting stream networks.

Hydro-enforcement did not improve the results of extracting the headwaters stream networks. Hydro-enforcement did, however, generate more direct drainages in the lake areas, indicating that flatter areas or areas prone to flooding, will be aided by breakline data. The breakline data available from the NC State Floodplain Mapping Program are better aligned with known streams and waterbodies than the available NHD stream centerlines (including the high-resolution data). Because hydro-enforcement will dictate the drainage for the higher-order streams, it is critical that the breakline data used for hydro-enforcement are of the highest quality, or that hydro-enforcement is not used.

Overall, the best results were obtained by using a hydro-enforced TIN for generating the base DEM and extracting the stream network from this base DEM using ArcHydro and the AGREE algorithm. This automated method produced a better headwaters stream map than currently available stream maps, including the manually derived Wake County hydrography lines.

The research questions and the findings associated with these questions are presented below:
What is the accuracy of the LIDAR surface elevation points? The vertical accuracy of the LIDAR elevation data was within .8 ft. (24 cm) as measured by the RMSE and using the 95 percentile technique (5% of the outlier points removed based on the “sigma-3” rule). This accuracy is particularly noteworthy as the published value of 25 cm is based on different measurement requirements (this study did not require the measurement to be taken on uniformly sloping terrain and included only one landcover type).

When using densely spaced LIDAR data, does the interpolation method used to create a DEM make a difference in the accuracy of the DEM? The interpolation methods compared (natural neighbor, IDW, spline, and kriging) produced similar results as compared to the survey data. Two methods, however, were superior in modeling the channel terrain—natural neighbor (RMSE = 1.6 ft.) and spline (RMSE = 1.4 ft.). Kriging required expertise to fit the parametric model to the semivariogram and substantially more computational power and time, without improving the results. It should be noted, however, that the study area contained good LIDAR coverage. More research needs to be done to compare interpolation methods in areas void of LIDAR data.

At what resolution do you begin to lose channel definition of headwaters streams, thereby affecting the production of headwaters stream maps? The stream networks extracted from the various DEM resolutions revealed that the 20 ft. resolution DEM is the most optimal for extracting stream networks. Very little difference was noted between the stream networks extracted from the 10 ft. and 20 ft. resolution DEMs. The stream networks extracted from the 60 ft. and 90 ft. resolution DEMs exhibited an unnatural straight-line pattern and a loss of first order streams as compared to the drainage derived from the 20 ft. resolution DEM.
Does using breakline data to hydro-enforce TINs and DEMs make a difference in the production of headwaters stream maps? There were no substantial differences in the headwaters stream networks produced by the various stream extraction methods. The headwaters stream networks were substantially the same regardless of the interpolation method used to generate the DEM (natural neighbor, IDW and regularized spline), or whether the DEM was hydro-enforced. The most substantial differences in the stream extraction results were in the higher-order streams and were produced by; (1) the use of the AGREE algorithm to overlay the known stream centerlines (breaklines) onto the DEM, and (2) the accuracy of the breaklines used for hydro-enforcement. The best stream map was produced by the DEM generated using the hydro-enforced TIN and the AGREE algorithm.

For the crystalline felsic, fluvial eroded terrain of the Piedmont Ecoregion, this automated approach produced a better headwaters stream map than currently available stream maps, including the NHD high and medium resolution hydrography lines available for the State of North Carolina. In particular, the stream maps generated for the study area better depicted the location of the headwaters streams than the manually derived Wake County hydrography lines. Furthermore, the LIDAR-derived topographic maps generated as part of the stream extraction process were substantially more accurate than currently available USGS 7.5 minute DEMs. These improved topographic and stream maps will enable decision-makers to better assess and mitigate nonpoint source pollution problems, improve state planning activities such as road construction and industrial recruitment efforts, and aid the development and management of riparian buffer regulations and ordinances.
7.0 References Cited


Environmental Modeling Research Laboratory (EMRL) of Brigham Young University, WMS v7.0 Tutorials, Provo, Utah, 2002.


National Geodetic Survey (NGS), NGS Data Sheet for Emorys and Ridge monuments, [http://www.ngs.noaa.gov/cgi-bin/ds_county.prl](http://www.ngs.noaa.gov/cgi-bin/ds_county.prl), 1985 (document undated, date of last recorded entry)

National Geodetic Survey (NGS), Digital Survey DATA text file, [http://www.ngs.noaa.gov](http://www.ngs.noaa.gov), 2002


Poiker, T. K., Unit 39, NCGIA Core curriculum in geographic information science. Simon Fraser University, 1990.


US Army Core of Engineers, Control Station Data Sheet for FN-49, North Carolina Area Real Estate Office, Raleigh, North Carolina, 2002


Figure 3.1: Develop survey plan and obtain field data for validating high-accuracy LIDAR elevation data, modeled elevation, and stream extraction models.
Figure 3.2: Compare LIDAR elevation data and interpolation methods for creating topographic maps (20ft resolution).
Figure 3.3: Evaluate the effect of scaling on topographic map accuracy and stream delineation (90ft, 60ft, 20ft, 10ft resolutions).
Figure 3.4: Assess the impact of using breakline data to extract stream networks (20ft. resolution).
Figure 4.1: Study area near Falls Lake, North Carolina. Control benchmarks set along ridge are shown in red. The study catchment is outlined in white.
Figure 4.2: Close-up of study catchment (outlined in white). Red points are control benchmarks; blue points are transect surveyed across the catchment, and; yellow points are mapping-grade GPS taken along the headwater stream channel.
Figure 5.1.1(a): Elevation of survey points (blue) and LIDAR TIN (red) for transect taken across the headwater stream catchment. The LIDAR points align well with the survey points except at the channel bottom.

Figure 5.1.1(b): Close-up of channel area. Survey points (blue) show small ephemeral and channel bottom missed by LIDAR (as predicted by TIN) (red).

<table>
<thead>
<tr>
<th>Data Description</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIDAR</td>
<td>1.2'</td>
</tr>
</tbody>
</table>

76
Figure 5.2.1: Elevation of survey points (blue), natural neighbor [TIN-to-grid] (red), and spline [regularized] (green) for transect taken across the headwater stream catchment.

<table>
<thead>
<tr>
<th>Data Description</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey</td>
<td>1.2'</td>
</tr>
<tr>
<td>Kriging</td>
<td>2.3'</td>
</tr>
<tr>
<td>Natural Neighbor</td>
<td>1.6'</td>
</tr>
</tbody>
</table>

Figure 5.2.2: Elevation of survey points (blue), kriging [point-to-grid] (red), and IDW [variable] (green) for transect taken across the headwater stream catchment.

<table>
<thead>
<tr>
<th>Data Description</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey</td>
<td>1.2'</td>
</tr>
<tr>
<td>IDW</td>
<td>2.2'</td>
</tr>
<tr>
<td>Kriging</td>
<td>2.3'</td>
</tr>
</tbody>
</table>
Figure 5.3.1: Comparison of 10 ft. and 20 ft. resolution DEMs. Shown here are the elevation values for survey (blue), spline [regularized – 10 ft.] (red), and spline [regularized – 20 ft.] green). The two DEMs have minor differences.

<table>
<thead>
<tr>
<th>Data Description</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIDAR</td>
<td>1.2'</td>
</tr>
<tr>
<td>Spline – 10 ft.</td>
<td>1.3'</td>
</tr>
<tr>
<td>Spline – 20 ft.</td>
<td>1.4'</td>
</tr>
</tbody>
</table>
Figure 5.3.2: Comparison of 60 ft. and 90 ft. resolution DEMs. Elevation of survey (blue), spline [regularized – 60ft.] (red), and spline [regularized – 90ft.] (green). At 60 ft. resolution, the 2 ft. channel becomes a 50 ft. depression. At 90 ft. resolution, the entire drainage is lost. 

<table>
<thead>
<tr>
<th>Data Description</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline (20 ft. resolution)</td>
<td>1.4'</td>
</tr>
<tr>
<td>Spline (60 ft. resolution)</td>
<td>5.6'</td>
</tr>
<tr>
<td>Spline (90 ft. resolution)</td>
<td>7.0'</td>
</tr>
</tbody>
</table>
Figure 5.3.3: Comparison of stream drainages extracted from 10 ft. (red) and 20 ft. (blue) resolution DEMs. Very little difference exists between the two networks.
Figure 5.3.4: Stream drainage extracted from 60 ft. (blue) resolution DEM overlain on breaklines (blue). Note the loss of first order streams as compared to Figure 5.3.3 and the linear bias in the “drawing” of the streams highlighted by the red ovals.
Figure 5.3.5: North Carolina State Floodplain Program breaklines (red) overlain on 1999 Wake County orthophotography. Note good alignment of breaklines with natural drainages
Figure 5.3.6: NHD [high resolution] (red) overlain on breaklines (green) and Wake County 1999 orthophotography. Alignment problems are highlighted with black ovals.
Figure 5.3.7 Shows TIN generated with breaklines
Figure 5.3.8: Stream network extracted from DEM generated from hydro-enforced TIN [TINgrid with breaks/with AGREE] (blue) compared to DEM generated directly from LiDAR using spline [spline-reg with AGREE] (brick). Tags along lake centerline occur when the hydro-enforced TIN is not used. Otherwise, the methods produce the same result.
Figure 5.3.9: Difference between stream networks extracted by overlaying breaklines using AGREE (blue) and not overlaying breaklines (red). Note the confused drainage lines in the lake area, indicating that using breaklines may make a difference in areas that are flat and prone to flooding.
Figure 5.3.10: “Best” modeled drainage (blue) produced by using the hydro-enforced TIN to generate a DEM (natural neighbor) and then extracting the stream network using the AGREE algorithm (Table 4.4.1: TINgrid with breaks/with AGREE). Modeled drainage is overlain on the manually derived Wake County hydrological lines (green) for comparison.
Figure 5.3.11: Modeled drainage (blue) does better job than Wake County hydrography lines (red) in following actual stream for the study area (yellow GPS points).

Figure 5.3.12: Modeled drainage overlain on USGS 1:24,000 Topographic DRG shows contour “v’s” depicting drainage are off by an average of 110 feet for differences measured.
Figure 5.3.13: Comparison of stream lengths for stream extraction methods as compared to other sources. Note the added stream length beyond the breakline file generated by the stream extraction method.