

Generating Stream Maps Using LiDAR Derived Digital Elevation Models and 10-m
USGS DEM

Adam Mouton

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2005

Program Authorized to Offer Degree:
Forest Resources

ACKNOWLEDGEMENTS

I am grateful to Peter Schiess for giving me the opportunity to take on this project and work through it as desired and Finn Krogstad for his ability to handle a bombardment of questions. Thank you to David Montgomery and Steven Burges for advising me on the ways of hydrology and then some; Luke Rogers and Phil Hurvitz for early GIS advice; Hans-Erik Andersen for LiDAR pre-processing; Bob McGaughey for various advice; David Tarboton and Theodore Endreny for question regarding resolution and modeling; and Julie Forcier for grammatical help. Capstone 2005 which is composed of Adam Baines, Lou Beck, Ben Carlson, Mark Williams, Sara Wilson, Edwin Wong, and Amy Hawk for there assistance during field collection. The Washington State DNR provided the financial support for this study.

University of Washington

Abstract

Generating Stream Maps Using LiDAR Derived Digital Elevation Models and 10-m
USGS DEM

Adam Mouton

Chair of the Supervisory Committee:
Professor Peter Schiess
Forest Resources

The effects of digital elevation model (DEM) grid size for stream network predictions in the northwestern United States were examined to test the accuracy of high-resolution LiDAR (Light Detection And Ranging) digital elevation data. LiDAR elevation data were gridded at 2-, 6-, and 10-m scales and flow paths were predicted by four common routing algorithms known as D8, D-Infinity, Multiple Flow, and DEMON, D8 being the least sophisticated. These routing algorithms were also applied to a 10-m USGS DEM to compare LiDAR with the previously used data for hydrologic modeling. The analyses indicated that as topographic detail increased, all LiDAR-derived models delineated more streams and located streams in their topographically correct position when compared to a 10-m USGS DEM. Stream maps generated by either D8 or DEMON converged as the DEM resolution was increased. The data suggests that increased DEM resolution decreases the need for sophisticated models, reducing processing times required to create accurate stream locations and attributes.

LiDAR digital elevation data also improved the modeling of perennial stream heads and fish habitat potential in a direct comparison to a 10-m USGS DEM. Distances between stream heads predicted using a LiDAR dataset and field verified stream heads were significantly less than those predicted using a USGS dataset. This illustrates the potential use of LiDAR to accurately predict perennial flow in a given landscape. The ability to

locate fish barriers based on landscape gradient also improved with LiDAR data. A USGS dataset used to find fish barriers occasionally found barriers in places where none existed or vice versa. As LiDAR datasets become more available, automated creation of stream networks and their hydrologic features will become more feasible and the accuracy of the results will be much improved.

TABLE OF CONTENTS

List of Figures	iii
List of Tables	iv
1. INTRODUCTION	1
1.1 Overview.....	1
1.2 Previous Studies and Background Review	2
1.3 Current Stream Data	6
1.4 Research Objectives.....	8
2. METHODS / MODEL DEVELOPMENT	9
2.1 Site Description.....	9
2.2 Stream Model.....	11
2.3 Water-Typing Models.....	11
2.3.1 CMER Model.....	11
2.3.2 Gradient Model	13
2.4 PIP Method	13
2.5 Flow Direction Methods Utilized	16
2.6 Stream Channel Determination by Flow Accumulation.....	18
2.7 Generation of Flow Direction Networks.....	19
2.8 Current Hydro Layer Evaluation	19
2.8 Field Measurements	20
3. RESULTS	21
3.1 Flow Direction Comparison.....	21
3.2 Resolution Effects on Flow Direction.....	25
3.3 Assessing the Current Hydro Layer.....	28
3.4 Determining Perennial Streams	29
3.5 Determining Fish Stream Habitat	34
3.6 Weakness and Shortcomings	37
4. CONCLUSION.....	38

REFERENCES	39
APPENDIX A: LiDAR System Used.....	42
APPENDIX B: LiDAR Data Collection.....	44
APPENDIX C: GEN2ASC.EXE	48
APPENDIX D: TXT2TIN.AML	53
APPENDIX E: Virtual Deforestation and Aggregation AMLs	54
APPENDIX F: Programs Used For Flow Direction.....	59
APPENDIX G: Regression Correlation of Flow Direction Algorithms with Increased Resolution.....	61
APPENDIX H: D8 Comparison at 10-m Resolution of USGS and LiDAR Generated Data.....	69
APPENDIX I: Statistics for Binary Linear Regression	72

LIST OF FIGURES

Figure Number	Page
Figure 1: Definitions of concentrated and dispersed contributing area and specific catchment area	2
Figure 2: Hydro Layer Discrepancies	3
Figure 3: 10 Meter USGS Contours overlaid onto a LiDAR DEM.....	4
Figure 4: Stream layer produced from flow accumulation in GIS.....	5
Figure 5: Study Site	10
Figure 6: Points centered around the Stream head which indicates the status of perennial flow	15
Figure 7: Schematic of the DEM pixel aspect computation and flow angle mapping performed by the D8, MFD, Dinf, and DEMON algorithms....	18
Figure 8: Plot of catchment area for 2-m LiDAR DEM at Tahoma State Forest	22
Figure 9: Plot of catchment area for 10-m USGS DEM at Tahoma State Forest.....	23
Figure 10: USGS and LiDAR Generated 10-m, D8 Comparison.....	25
Figure 11: 2-m DEM without using culvert correction	26
Figure 12: 2-m DEM using culvert correct. Stream culverts are circle, Ditch culvert are triangle	27
Figure 13: 6-m DEM stream channel determination	27
Figure 14: LiDAR vs. DNR perennial streams.....	28
Figure 15: Stream head defined by the landscape	29
Figure 16: Spring identified as a stream head in the field	30
Figure 17: Stream heads within the study site	31
Figure 18: The change in distance field verified stream heads have from the modeled stream heads at a given resolution	33
Figure 19: Field Stream Head distance from various flow direction modeled stream.....	34
Figure 20: Longitudinal Profile of a selected creek.....	35
Figure 21: Stream Channel with predicted fish habitat	36

LIST OF TABLES

Table Number	Page
Table 1: Area Covered by Buffers	6
Table 2: Summary of the final logistic regression model coefficients, standard errors, significance of the coefficients, and 95% confidence intervals for the exponential of the coefficients	13
Table 3: Stream parameters identified and collected in the field.....	20
Table 4: Values associated with each stream feature type.....	20
Table 5: Relation between D8 and DEMON algorithms at various resolutions.....	23
Table 6: DNR Hydro Layer Potential Error.....	29
Table 7: Sub-Basins used in perennial head identification.....	30
Table 8: Summary of the final logistic regression model	31
Table 9: Predicted Fish-Bearing Streams within the Study Site using different techniques.	36

1. INTRODUCTION

1.1 Overview

Streams are one of the most valuable public resources in the State of Washington. They provide the habitat for fish, an important cultural as well ecological asset. Fish bearing streams and the related habitat also are typically found within forests, important for various aspects, such as stream quality and economic values. Ironically, there is a lack of accurate data in order to properly delineate streams. Long-term sustainable harvest volume calculations, feasible harvest settings and road location design at the landscape or watershed level are critically dependent on reliable stream data. In a small project near Forks (Schiess and Tryall, 2002), stream buffer areas based on official DNR data (Hydro layer) underestimated actual stream area by an average factor of two.

However, new mapping technology provides the potential of developing improved stream data from more detailed surface topology. LiDAR (Light Detection And Ranging) data which creates sub meter topography maps (Appendix A) is one technology that promises to provide increased resolution in digital surface detail compared with the typical 10 meter topographic maps and could lead to more precise and accurate maps of stream networks. Preliminary analyses showed that using LiDAR data located more actual stream channels and placed streams in their topographically correct position (Schiess and Tryall, 2002). This ability to generate accurate stream locations and physical attributes using LiDAR will allow long-term sustainable harvest volume calculations to be more reliable.

In order to properly map the physical extent of channels in a watershed, the difference between processes on hillslopes and in channels must be determined (Tarboton 2003). This difference becomes apparent when calculating how water collects on a landscape in a given dataset with flow direction of the water known.

In channels flow is concentrated. The drainage area, A , (in m^2) contributing to each point in a channel may be quantified. On hillslopes flow is dispersed. The "area" draining to a point is zero because the width of a flow path to a point disappears. On hillslopes flow and drainage area need to be characterized per unit width ($m^3/s/m = m^2/s$ for flow). The specific catchment area, a , is defined as the upslope drainage area per unit contour width, b , ($a = A/b$) (Moore 1991) and has units of length ($m^2/m = m$). (Tarboton 2003 p.1-2) Figure 1 illustrates these concepts.

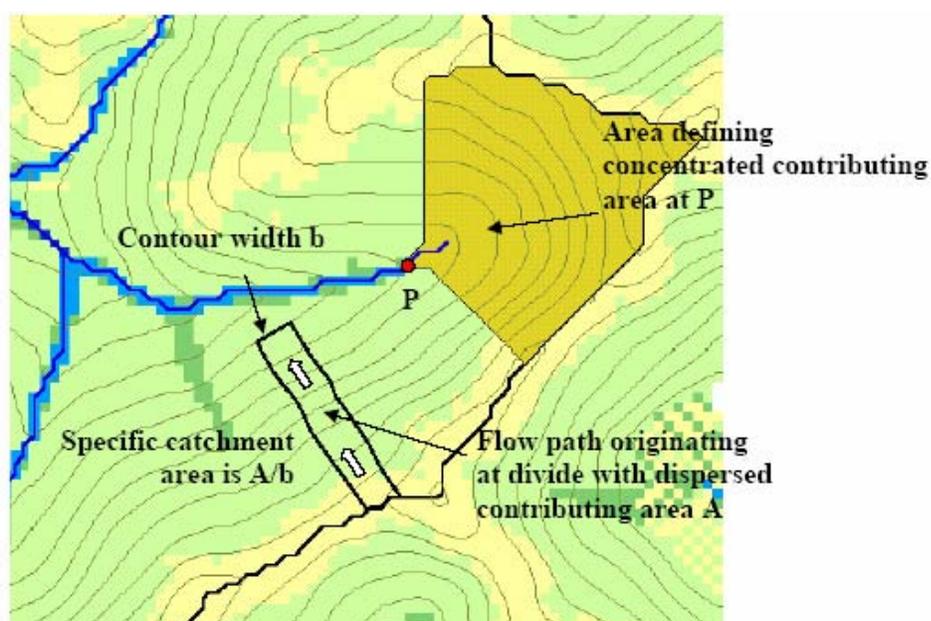


Figure 1. Definitions of concentrated and dispersed contributing area and specific catchment area (Tarboton 2003).

LiDAR can place these physical attributes in their topographically correct position. However, locating point 'P' (Figure 1), the perennial initiation point (PIP), becomes a challenge due to it being dependent on the catchment area which fluctuates based on geology, climate, precipitation, and other attributes.

1.2 Previous Studies and Background Review

Various hydrological models such as Simulator for Water Resources in Rural Basins (SWRRB), Environmental Policy Integrated Climate (EPIC), Groundwater Loading Effects of Agricultural Management Systems (GLEAMS), TR20, HEC-1, and HEC-2

have been used in modeling stream networks (Luijten 2000). The abundance of models demonstrates the importance given to modeling hydrologic features. During their development time, typical grid models used 5 – 90 meter DEM's. This leads to hydrologic maps not containing the full stream network, and at times, streams that are topographically incorrect.

Figure 2 demonstrates topographic error in the Washington DNR stream layer, in yellow, not extending completely up the channel. Overlaid is a 2m LiDAR-derived hillshade model that depicts the presumably correct stream courses. The DNR stream is also shifted by 300 feet to the east of the channel (Schiess and Tryall, 2002). Such discrepancies are not uncommon. This example illustrates possible inaccuracies when stream locations are determined using 7.5-minute topomaps, orthophotos, and some field verification. This was demonstrated in other projects as well and usually is recognized by field staff and planning staff as a critical issue in developing reliable forest operations designs.

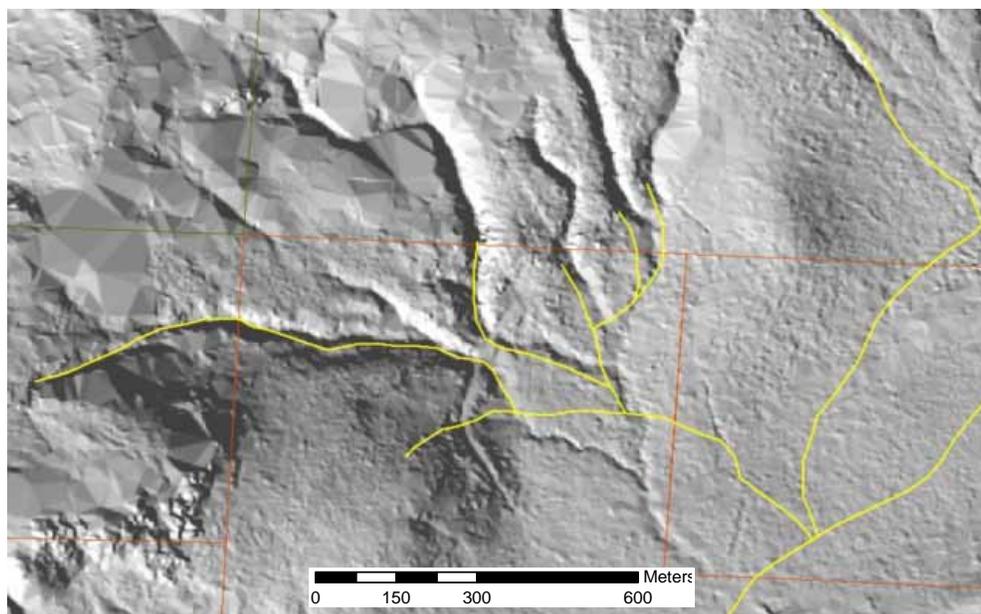


Figure 2. DNR stream data (yellow lines) overlying a 2m LiDAR-derived hillshade model that depicts the presumably correct stream courses. Note the discrepancies between LiDAR-derived hydrographic features and 7.5-minute-derived hydro data residing in the official data layer (Schiess and Tryall, 2002).

Figure 3 shows the contours generated from a standard 10-m DEM overlaid on a 2m LiDAR-derived hillshade model with slope classes from the LiDAR derived DEM. Downhill is toward the upper right. LiDAR topography provided a realistic and detailed topography. Photogrammetrically produced contour lines captured the general shape of the landscape; however, complex features such as incised streams, draws, abandoned road beds and sharp ridges were not recognized (Schiess and Krogstad, 2003). The contour lines also do not follow the stream channel accurately.

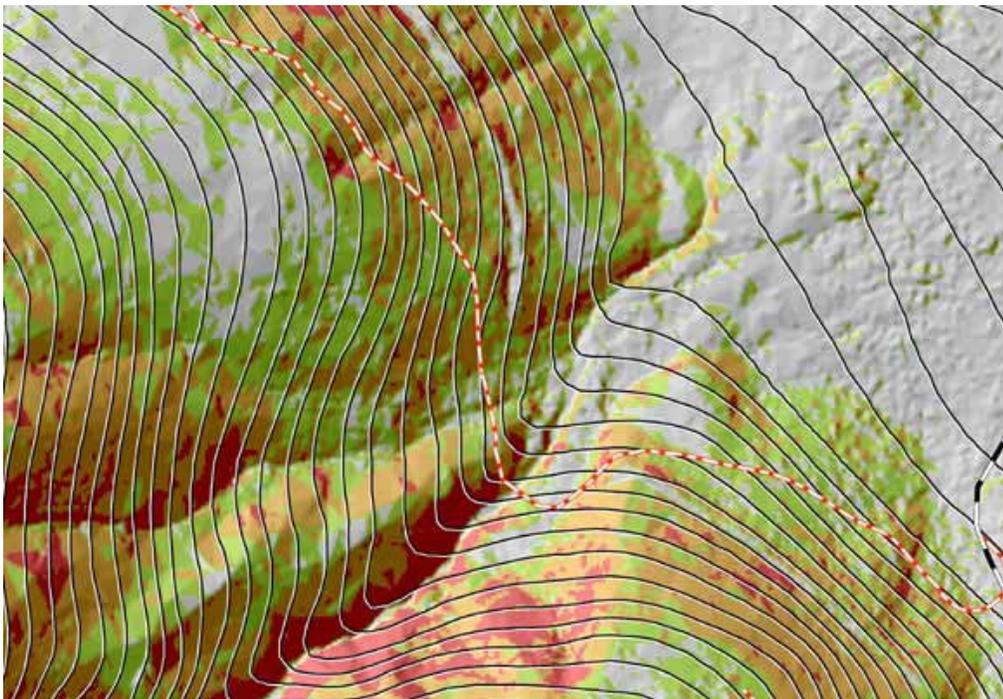


Figure 3. A Lidar-hillshade derived from 2m grids versus the contours, derived from DNR's 1:4800 photogrammetrically derived maps with roads. Downhill is toward the upper right. LiDAR topography provided a realistic and detailed topography. Photogrammetrically produced contour lines captured the general shape of the landscape. However, those contour lines miss the topographically and regulatory important stream depression as indicated by the LiDAR hillshade model (Schiess and Krogstad, 2003).

The advantage of LiDAR digital data over conventional photogrammetry is improved mapping in obscured areas. A LiDAR bare ground surface model containing only elevations can be obtained after filtering out the trees and buildings in the dataset. The digital data then can be used in a variety of ways including: digital terrain model for use

in generating contours, 3D terrain views, fault locations, steep slopes, critical areas, and stream and drainage basin delineation (North Carolina 2003).

One attempt to use LiDAR data to generate stream channels at the College of Forest Resources, University of Washington was on the South Tye Planning Study in 2002 (Schiess and Tryall, 2002). The stream layer was produced by the “flowaccumulation” command in GRID and a uniform buffer was added (Figure 4). The stream layer could be adjusted by changing the contributing cells to the stream, which made it possible to duplicate conditions observed in the field. The contributing cell size was adjusted to a slightly higher level in order to include areas that may not have contained water at the time. It should be noted that while the GIS method of “flowaccumulation” puts streams in their expected channels, it can both over- and underestimate the stream lengths (Schiess and Tryall, 2002). This is because a uniform catchment size is defined on all stream basins when catchment size could vary from basin to basin which causes inaccuracy.

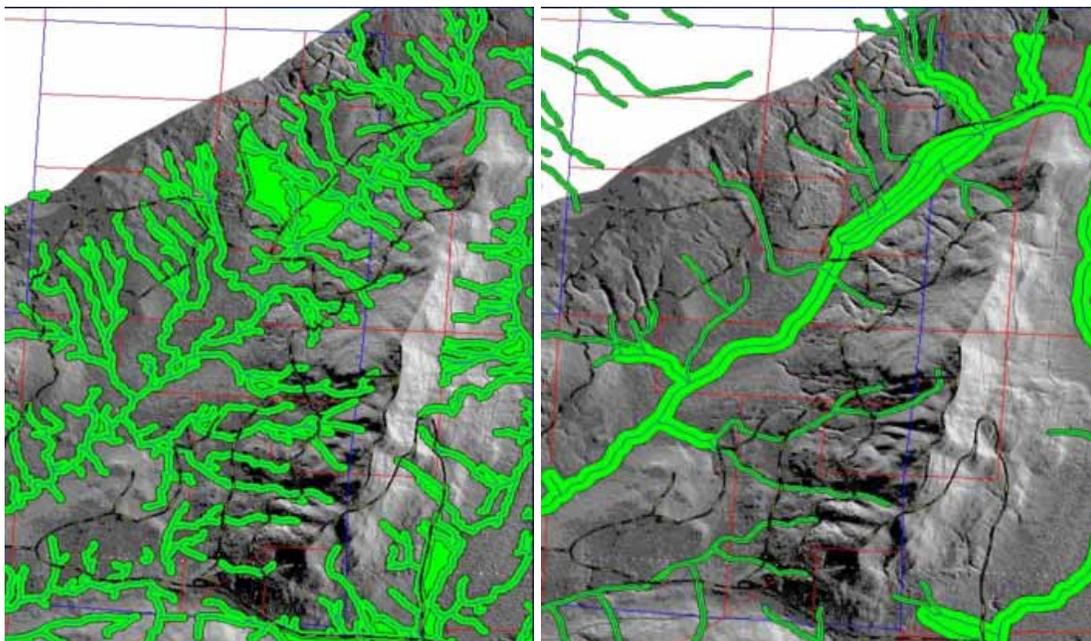


Figure 4. Stream layer produced from flow accumulation model in ArcGIS with 75-ft buffer using LiDAR at left compared to DNR hydro layer with buffered widths scaled based on stream type (Schiess and Tryall, 2002).

There were many discrepancies in comparing the LiDAR “flowaccumulation” streams with a DNR-Hydro layer, which is based partly on a 7.5 minute topographic generated streams in the South Tyee Planning Study (Schiess and Tryall, 2002). The 7.5 minute streams were buffered with widths based on stream type and the LiDAR streams were buffered at an average of 23m. Table 1 shows that there is a 182-ha difference between the two buffer representations. No thorough field verification was conducted, however.

Table 1. Area Covered by Buffers in South Tyee (Schiess and Tryall, 2002).

Stream Type	Buffer Area (ha)	% of Study Area
7.5 Minute Streams	142	21
LIDAR Streams	324	47

With better field reconnaissance and appropriate buffer widths a more accurate stream layer could be produced. However, the LiDAR stream data provided a better input for the preliminary planning process than the 7.5 minute stream data (Schiess and Tryall, 2002).

1.3 Current Stream Data

The current stream data was created by the DNR and is accessible through their online database. The stream data represents an integrated network coverage (polygons and lines) that contains data on water bodies (open waters, lakes, etc.) and watercourses (rivers, streams, canals, etc.) (Hydro metadata). The data was produced using orthographic photos from 1974, topographic maps, and field observations. On March 1, 2005 a new Water Type Attribution was completed. The primary purpose of the DNR hydro layer was to aid in the application of timber harvest and other forest practices regulations and activities by the Washington Department of Natural Resources (DNR). Other uses include cartography and analysis where hydrographic data is required.

The Water Type process occurred during a time of significant and rapid improvement in technical information and software tools. As a result of the extensive fish surveys being performed, abundant field survey information was available for many areas of the state. Advances in GIS technology provided opportunities to evaluate resource

protection and economic performance of alternative water typing systems across large geographic areas. Digital Elevation Models (DEM) produced by the U. S. Geologic Survey became widely available, allowing for consistent and reliable characterization of the physical landscape. For the first time since the implementation of forest practice regulations governing fish-bearing water bodies in the 1970s, the tools and data were available to develop and assess a data-driven classification system for use across the entire state. (Conrad et al., 2003 p.4)

The physical attributes for the Water Typing model were based on a USGS 10-DEM. Using this DEM, physical barriers such as waterfalls and downstream gradient could be overlooked due to the low resolution. Furthermore stream channels predicted using aerial photos under and over estimate stream locations due to visibility. At times, the channels were topographically off from the actual location of those channels (Schiess and Tryall, 2002).

The current DNR hydro layer has two coding systems, type code and fish/non-fish code. Type code is describes as:

- type 1, 2, and 3 ----- Fish bearing waters
- type 4 and 5----- Non-fish bearing waters
- type 9----- Untyped, unknown

Types 1-4 are considered perennial and type 5 and 9 are seasonal. The second code either describes streams as fish-bearing or non-fish bearing waters. This code was derived from the Cooperative Monitoring, Evaluation, and Research group (CMER) using the CMER Model as described in the Method / Model Development section. The same channel network is used for both code systems.

1.4 Research Objectives

The goal of this project is to determine if LiDAR would improve stream network classification. Therefore the following questions needed to be answered:

- Does an increase in resolution improve stream channel determination?

- Can stream types be determined more accurately using LiDAR datasets?
- Can a new algorithm be developed for identifying perennial streams?

To verify that the increased LiDAR resolution improves stream modeling, different hydrologic models were tested using a 10 meter USGS and several LiDAR DEMs at various resolutions. D8, D-Infinity, Multiple Flow, and DEMON were the model algorithms used (refer to section 2.5 Flow Direction Methods Utilized). Once the models were used and data was generated from the model, field verification was carried out to verify the accuracy of the predicted stream channels.

2. METHODS / MODEL DEVELOPMENT

Flow direction algorithms for locating stream channels were used on various resolutions and correlated with field data. This was completed to establish which flow direction technique worked best with LiDAR data. For stream typing, the Cooperative Monitoring, Evaluation, and Research group (CMER) from the Washington State Department of Natural Resources (DNR) has established a model which predicts which streams are inhabited by fish and which do not contain fish. This model was compared to a gradient model approach which used a LiDAR DEM.

Three models needed to be developed in order to decide if resolution has an effect on stream channel determinations and if stream types could be determined more accurately by LiDAR. The first model is the generation of the stream network from LiDAR DEM. The second is a water-typing model to determine the end of fish point (EOFP) from the generated stream network and the third is the perennial initiation point (PIP) model.

2.1 Site Description

The North fork of the Mineral Creek Watershed in the Mt. Tahoma State Forest is an area of approximately 3600-ha of forested terrain on steep topography with slopes up to 80% (Figure 5). It is located near Ashford WA, contained within T14N, R6E and T13N, R6E. This site was chosen because the Forest Engineering (FE) capstone project was located there and could provide logistical support as well as utilizing initial findings in the development of a forest transportation strategy which was critically dependent on a reliable stream location depiction. Digital datasets of the existing hydrology, cross drains, roads, fish barriers, soils, and high resolution digital elevation models based on LiDAR were obtained from WA DNR. The LiDAR dataset was flown in February 2003 and processed by the University of Washington with help from Hans-Erik Andersen and Matt Walsh (Appendix A). The DEM's that were processed from the LiDAR at 2-m grid cell size were used for forest transportation designs as part of the FE Capstone projects in 2003 and 2005 (Schiess and Tryall, 2003; Schiess and Mouton, 2005).

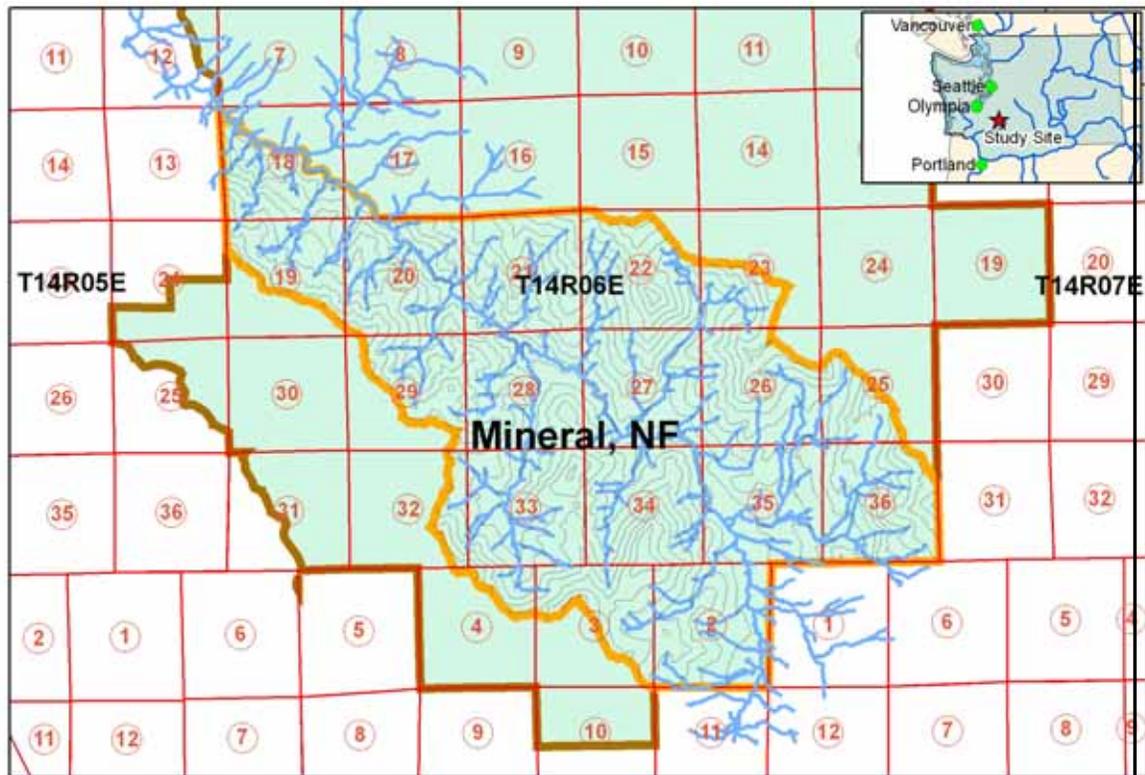


Figure 5. The North fork of the Mineral Creek Watershed in the Mt. Tahoma State Forest within the orange boundary is an area of approximately 3600-ha. (60-m contour lines)

The mean annual rainfall in this area ranges from 2007 to 2210 mm (Daly et al. 2000) with an altitude range of 500 to 1600 m. Forest cover is dominated by Douglas Fir (*Pseudotsuga menziesii*) with Western Larch (*Larix occidentalis*), Red Alder (*Alnus rubra*), Big Leaf Maple (*Acer macrophyllum*), Western Hemlock (*Tsuga heterophylla*), White Fir (*Abies concolor*), and Black Cottonwood (*Populus trichocarpa*) throughout. The majority of the region's soils belong to the Bellicum, Cattcreek, and Cotteral soil series (Soil Survey Staff, 1998). The upper part of the profile has a cindery texture from the pumice and volcanic ash aerially deposited from Mt. St. Helens. The lower part of the profile formed in colluvium, alluvium or glacial till from andesite with a mixture of pumice and volcanic ash. In general, these soils have low fertility and water-holding capacity and often occur on unstable slopes. The geology is categorized by Oligocene-Eocene (OEvba) and Oligocene (Ovc(oh)) defined as basaltic andesite flows and volcanoclastic deposits or rocks.

2.2 Stream Model

The type of streams that were modeled included all segments of natural waters within the bankfull widths of defined channels which are either perennial streams (waters that do not go dry any time of a year of normal rainfall) or were physically connected by an above-ground channel system to downstream waters. In extracting networks from DEM's, Tarboton et al. (1991) suggest that the network extraction should have properties traditionally ascribed to channel networks and have as high resolution as possible. A LiDAR-generated DEM provides this high resolution and the physical properties, such as channel depth and slope, associated with a stream network.

2.3 Water-Typing Models

Two preexisting logistic regression models were used to identify potential fish habitat. The CMER Model takes into account several physical attributes while the Gradient Model focuses on the gradient of the landscape. LiDAR DEM data was used to generate the attributes needed to apply these models.

2.3.1 CMER Model The CMER Model was used to identify potential fish habitat. Based on previous research, this habitat-based, water-typing model was developed using logistic regression analysis and GIS data which incorporated the results of field surveys (Conrad et al., 2003).

The fish absent | fish present (FAFP) data used to estimate the logistic regression models were generated from 4,052 end-of-fish points (EOFP) placed on a Washington Dept. of Natural Resources (DNR) GIS hydrologic layer. Each EOFP was based on a field survey which followed specific protocols to identify a location on the stream that was designated as either last fish or last fish habitat. Potential EOFPs were submitted to DNR for error checking and initial screening (Conrad et al., 2003). After approval by DNR, the EOFP was transferred from the DNR hydrologic layer to a 10m DEM-generated stream point network. An automated procedure was then used to classify points upstream of the EOFP

as fish absent points and points downstream of the EOFP as fish present points. There were four physical attributes associated with each point on the DEM network:

1. Basin size (number of acres in surrounding basin that drain through a point),
2. Elevation in feet (based on 10-m DEM network),
3. Downstream gradient which is the average gradient measured over 100-m downstream of the point (calculated from 10-m DEM network elevation information), and
4. Precipitation in inches (GIS derived estimate of average annual precipitation at the point based on Daly et al. [1998]).

These four physical attributes associated with each EOFP point were the variables available for the logistic regression model building process.

Equation 1 is the response function where π is the estimated probability of fish presence.

$$(1) \quad \pi = \frac{e^{(\beta'\chi)}}{1 + e^{(\beta'\chi)}}$$

Equation 2 is the linear model where the variables are described in Table 2 (Conrad et al., 2003).

$$(2) \quad \beta'\chi = -7.717073 + (0.020166 * (\text{PRECIP})) + (3.793994 * (\text{Log10}(\text{BASIZE}))) - (0.062949 * (\text{DNGRD})) - (0.110926 * (\text{ELEV} / 100))$$

The CMER Water-Typing Model was applied to the LiDAR DEM in the same manner it was applied to the 10m DEM. Downstream gradient (DNGRD), Elevation (ELEV), and Basin Size (BASIZE) were determined by LiDAR using the LiDAR-derived stream network and elevation model. Table 2 provides the results of the logistic regression model (Conrad et al., 2003).

Table 2. Summary of the final logistic regression model coefficients, standard errors, significance of the coefficients, and 95% confidence intervals for the exponential of the coefficients. (Conrad et al., 2003)

Coefficient	Estimate	Standard Error	Significance	e^B	95% Confidence Interval	
Constant	-7.717073	0.25192	<0.001			
\log_{10} (BASIZE)	3.793994	0.09625	<0.001	44.434	36.794	53.657
ELEV/100	-0.110926	0.00687	<0.001	0.895	0.883	0.907
DNGRD	-0.062949	0.00656	<0.001	0.939	0.927	0.951
PRECIP	0.020166	0.00191	<0.001	1.020	1.017	1.024

2.3.2 Gradient Model This model's objective was to identify the physical constraints and stream characteristics at the upstream limits of trout distribution (Latterell et al., 2003). Logistic regression was used to model the likelihood of trout presence in a 100-m stream reach as a function of physical stream attributes using sites described in Latterell et al. (2003) sites. The regression provided a probabilistic prediction of trout presence because the dependent variable was binomial (trout presence or absence). Further, this technique does not assume normality, equal variances, or a linear response. Equation (1) from above is the response function used for this model with $\beta'\chi$ as the linear model, which is

$$(3) \quad \beta'\chi = (DNGRD * B1) + B0$$

where B1 is the Gradient Coefficient at -0.209 and B0 is the Model Constant at 2.765. Logistic regression calculates the probability of success identified as (i.e., trout presence, $\pi \geq 0.50$) over the probability of failure (i.e., trout absence, $\pi < 0.50$).

2.4 PIP Model

The perennial initiation point (PIP) is the point where perennial flow begins on a Type 4 Water. Type 4 Water means all segments of natural waters within the bankfull width of defined channels that are perennial non-fish habitat streams. The model for PIP was developed in a GIS framework, whereby the measuring and recording of geomorphic stream characteristics were done by GIS at grid-center points along a LiDAR digital elevation model (DEM) generated stream network.

Perennial initiation point standard is defined in WAC 222-16-030(3) and 222-16-031(4) of the Washington State Register. For western Washington sites not in any coastal zone, Type 4 waters begin at a point along the channel where the contributing basin area is at least 21-ha.

The PIP Model was based on the techniques described in Conrad et al., (2003) except the logistic regression was used to identifying where perennial flow begins and where it ends instead of fish present. Stream head data locations were collected in the field for 5 sub-basins within a selected site. Points were generated 15 meters upstream and downstream of the stream head in GIS (Figure 6) for clear perennial definition and the following physical attributes were associated with the points to determine which influenced perennial flow:

1. Basin size – Using SAGA (System for Automated Geoscientific Analyses, a GIS system) (Appendix F) the algorithms described in the Stream Model section were utilized.
2. Downstream gradient - Using the 2-m LiDAR DEM, focalmin was used in ArcMap GIS 9.0 Raster Calculator using 100 meter mean. The 2-m DEM was then subtracted by the focalmin output and divided by 100 meters.

$$DG = \text{float}([2\text{m DEM}] - \text{focalmin}([2\text{m DEM}], \text{circle}, 50) / 328 * 100\%$$

50 is 100 meters in mapping units for 2m grid cell size and 328 is 100 meters in ft.

3. Forest Density - A spatial tree list was created using LiDAR by Hans-Erik Andersen. The table in the list was comprised of height and diameter of the tree. Crown radius was then calculated for each tree using:

$$CR = (H + .223) / 4.4$$

where CR is crown radius and H is height (Oladi 2001). A crown cover layer was created by buffering each point using the size of the crown radius and converting the resulting polygon coverage to a grid. Individual tree crown density was determined by taking an 11-m buffer around the tree point and dividing the area of the crown within the buffer by the area of the buffer.

4. Slope (calculated from 2m LiDAR DEM network elevation information in percent with ArcMap (%slope = 100 * Tan (Δy/Δx))),
5. Elevation in feet (LiDAR bare earth DEM),
6. Precipitation in inches (GIS derived annual precipitation at the point based on Daly et al. [2000]), and
7. Site Class (Downloaded from Washington DNR website)

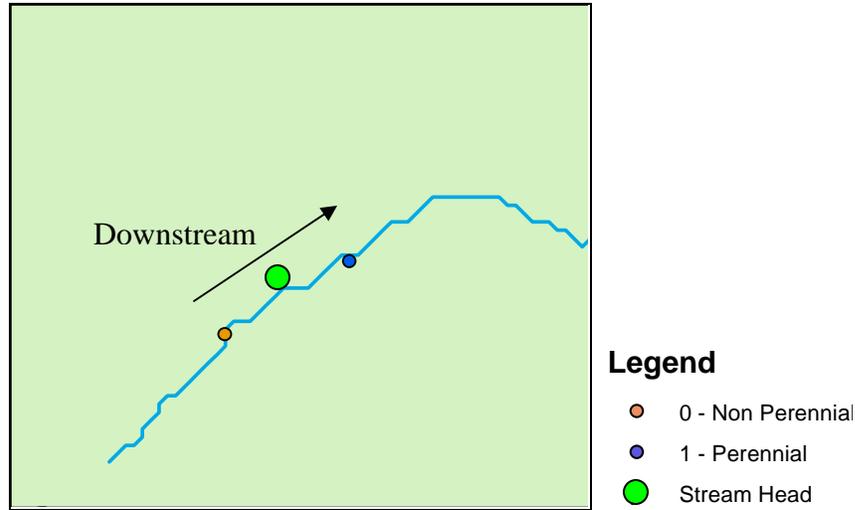


Figure 6. Plan view at a stream channel showing points generated 15 meters upstream and downstream of the stream head with physical attributes associated with the head. This is to produce a binomial linear regression model identifying perennial and non-perennial parts of a stream network.

The points upstream represented non-perennial flow as the points downstream represented perennial flow. Using binary logistic regression and setting 0 for non-perennial and 1 for perennial, the physical attributes associated with each point were used to develop Equation 1 from above and equation 4.

$$(4) \beta' \chi = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip}$$

where

$\beta' \chi$ is the estimated linear equation

x_{ij} is the j^{th} predictor for the i^{th} case of the physical attributes

b_j is the j^{th} coefficient for the physical attributes

p is the number of predictors.

After the model equation was applied and the values less than 0.5 (non-perennial, $\pi < 0.50$) were removed in Arc, regiongroup in ArcTools was applied to find the continuous parts of the network. The grid was then converted to an .asc file and imported into SAGA (Appendix F). Channel Network in SAGA was applied and the vector linear stream network was created.

After applying the model equation in Arc, probability of non-perennial was adjusted from 0.5 to around 0.97 for a conservative approach. This was determined by looking at the histogram in the classification class and moving the class line to the highest possible number without removing the stream network. Regiongroup in ArcTools was then applied to find the continuous parts of the network.

2.5 Flow Direction Methods Utilized

Four flow direction methods were utilized for this project. This includes the D8, D_{∞} , Multiple Flow Direction (MFD), and DEMON algorithms. The D_{∞} , MFD, and DEMON algorithms were utilized because the D8 algorithm has two major restrictions:

(1) flow which originates over a two-dimensional pixel is treated as a point source (non-dimensional) and is projected downslope by a line (one dimensional) (Moore and Grayson, 1991), and (2) the flow direction in each pixel is restricted to eight possibilities. (Costa-Cabral and Burges, 1994 p.1)

Spatial processing has a limited number of raster-based procedures. Collectively, these raster-based procedures implemented in ARC/INFO utilize the single flow direction (SFD) algorithm (O'Callaghan and Mark 1984). It calculates flow direction as the steepest slope direction, which is determined by the Maximum Downward Gradient (MDG) (Figure 5). This SFD algorithm, also known as the Direction 8 or D-8, is widely used on DEM data analysis and GIS software (e.g., the "Flowdirection" function in ARC/INFO GRID) (Jenson and Domingue 1988). The D-8 raster procedures form the basis for describing and modeling water flow through a digital elevation data set. Using the D-8 directional information associated with each cell, a network of flow from one cell

to the next is represented. Using the flow direction information, the number of cells that flow into any cell is tracked and assigned to the cell.

The D_{∞} algorithm, also known as Biflow Direction (BFD), was proposed by Tarboton (1997). In this algorithm the 3 X 3 cell window is divided into 8 triangular facets. The slope direction and magnitude of each facet are compared. The steepest downward direction is chosen and divided into two directions along the edges forming that facet (Figure 7). The proportion of flow along each edge is inversely proportional to the angle between the steepest downward directions and the edge. Therefore at most two flow directions can be assigned to each cell. The contour length is defined as the grid cell size (DEM resolution), and the slope is set to be the largest slope of 8 facets. (Tarboton 1997) (Pan et al., 2004 p.11)

Multiple Flow Direction (MFD) algorithm is the third method that was used for determining flow direction.

Quinn et al. (1991) first suggested this algorithm to improve representation of the convergence or divergence of flow. Wolock and McCabe (1995) showed how to implement this algorithm using ARC/INFO GRID functions. Unlike the SFD and BFD algorithms, the MFD algorithm, each flow direction is weighted by the downward elevation gradient (i.e., from the central cell to each of its 8 neighbors) multiplied by a “contour length” (Figure 7). There are two way so set the contour length: i.e. $(HR/2)$ and $(\sqrt{2}HR/4)$ (Quinn et al., 1991), or $0.6HR$ and $0.4HR$ (Wolock and McCabe 1995), for cardinal and diagonal flow directions, respectively, where HR is the horizontal resolution, (Pan et al., 2004 p. 2-3)

In DEMON (Digital Elevation MOdel Network) by Costa-Cabral and Burges (1994), flow is generated at each pixel (source pixel) and is routed down a stream tube until the edge of the DEM or a pit is encountered. The stream tubes are not constrained to coincide with the edges of cells and can expand and contract as they traverse divergent and convergent regions of the DEM surface. The stream tubes are

constructed from the points of intersections of a line drawn in the gradient direction (aspect) and a grid cell edge (Figure 7). The amount of flow, expressed as a fraction of the area of the source pixel, entering each pixel downstream of the source pixel is added to the flow accumulations value of the pixel. After flow has been generated on all pixels and its impact on each of the pixels has been added, the final flow accumulation value is the total upslope area contributing runoff to each pixel. (Wilson and Gallant, 2000 p.66)

Block Centered

12	11	9
9	10	8
11	7	5

Edge Centered

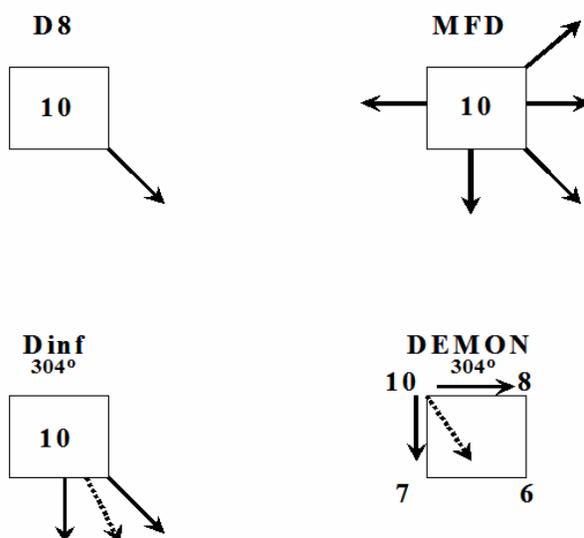
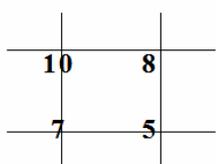


Figure 7. Schematic of the DEM pixel aspect computation and flow angle mapping performed by the D8, MFD, Dinf, and DEMON algorithms. Values on the left signify elevation. Solid arrows point in the direction that flow is mapped, and dotted lines correspond with the degree value above the pixel and indicate the pixel aspect for DEMON and Dinf. The distinctions between a block- and edge-centered DEM is illustrated (Endreny and Wood 2001)

2.6 Stream Channel Determination by Flow Accumulation

Once it is determined how water flows across a landscape, a flow accumulation value can be assigned to each cell showing how many cells are upstream from it. This simulates water as it accumulates going down hill to a stream channel. A user-determined accumulation cell threshold value identifies those cells that have concentrated flow and

those cells that do not have concentrated flow. This threshold value defines the catchment size required for flow to become a stream channel.

Montgomery and Dietrich (1992) found that landscape dissection into distinct valleys is limited by a threshold of channelization that sets a finite scale to the landscape. This threshold determines the number of sub-basins that exist in the model. A small threshold results in a large number of small basins whereas a large threshold results in a small number of large basins. The threshold is equal to the hillslope length / accumulation that is just shorter than that necessary to support a channel head (Montgomery and Dietrich 1992). A threshold-based approach is most appropriate for modeling channel head locations over shorter, geomorphic time scales (e.g. 10^2 - 10^3 years) than modeling valley development (e.g. 10^4 - 10^6 years) (Montgomery and E. Foufoula-Georgiou 1993).

2.7 Generation of Flow Direction Networks

Many programs were used to implement the different flow direction models on a LiDAR DEM. SAGA (System for Automated Geoscientific Analyses) was the main interface used while ArcMap, TauDEM (Terrain Analysis Using Digital Elevation Models), and TAS (Terrain Analysis System) were used to verify that SAGA was performing the flow direction models properly (Appendix F). The reason that SAGA was used instead of other programs was its ability to store grid files in memory. Memory storage allowed the data to be compressed saving computer memory to permit more demanding tasks such as computing flow direction algorithms. The LiDAR DEM had to be broken up into individual sub-basins for the other programs to use the data. Because of the computational demand of DEMON, the 2-m LiDAR DEM had to be divided into the basin specific for the study site for SAGA to utilize the dataset.

2.8 Current Hydro Layer Evaluation

Evaluation of the DNR hydro layer consisted of selecting two watersheds within the study site and confirming the DNR stream types correlated with what was out in the field.

When an error in correlation existed, the proper stream type was determined in the field to identify common problems within the DNR hydro layer.

2.9 Field Measurements

Field locations were sampled to test the accuracy of the current DNR hydro layer, various models used with the project, and evaluate the creation of a new stream model. Table 3 lists the features collected with Table 4 listing the associated attributes of those features.

Table 3. Stream parameters identified and collected in the field. The information was logged together with GPS coordinates.

Feature Type	Range within Feature Type	
	Width (ft)	Water Depth (ft)
Stream Head	1 - 10	0.1 - 3
Stream	1 - 7	0.2 - 5
Fish Blockage	1 - 15	0.5 - 6
No Channel	0	0
Channel - No Water	0	0
Seasonal	0 - 2	0 - 1
Other	0 - 1	0 - 1

Table 4. Fish presence or absence, % gradient and fish barrier type was logged with each stream feature type.

Fish Present	Possible Values	
	Gradient (%)	Fish Barrier
No	0 - 4	Culvert
Yes	5 - 9	Logs
	10 - 15	None
	16 - 50	Waterfall
		Rocks
		Road

Field planning was done by using LiDAR streams with the DNR hydro layer to locate stream features in the field. The above values were then determined and logged using a Trimble ProXRS GPS unit. The ranges for width and water depth for each stream feature were evaluated (Table 3) as well as the possible values (Table 4) that the features could have associated with it. Fish presence was assessed only by visual inspection of the stream. If a stream did not contain much water, contained many fish barriers, or went underground, it was assumed that no fish would be present. Field work was done April 20th through May 3rd 2005.

3. RESULTS

3.1 Flow Direction Comparison

The D8, Dinf, MFD, and DEMON predicted stream channel locations differently. The major difference in stream channel determination was in hillslope interpretation.

Differences between D8 and DEMON decreased as resolution increased. From visually inspecting Figure 8, a 1:1 relation between D8 and DEMON at the 2-m resolution was shown using regression correlation, when a stream is defined as a 5-ha (13.2 (ln ft²)) basin. The models do not begin to correlate until 30.4-ha (15.0 (ln ft²)) at the 10-m USGS resolution from Figure 9. As resolution increased, the spread of the points in those figures becomes more confined to the lower left corner indicating the correlation between D8 and DEMON flow direction models. Table 5 summarizes the relationship between D8 and DEMON algorithms at various resolutions. No correlation can be determined with any combination of the other flow direction models with respect to increased resolution (Appendix G). APPENDIX G provides the plots of catchment area at all resolutions with the flow direction models used.

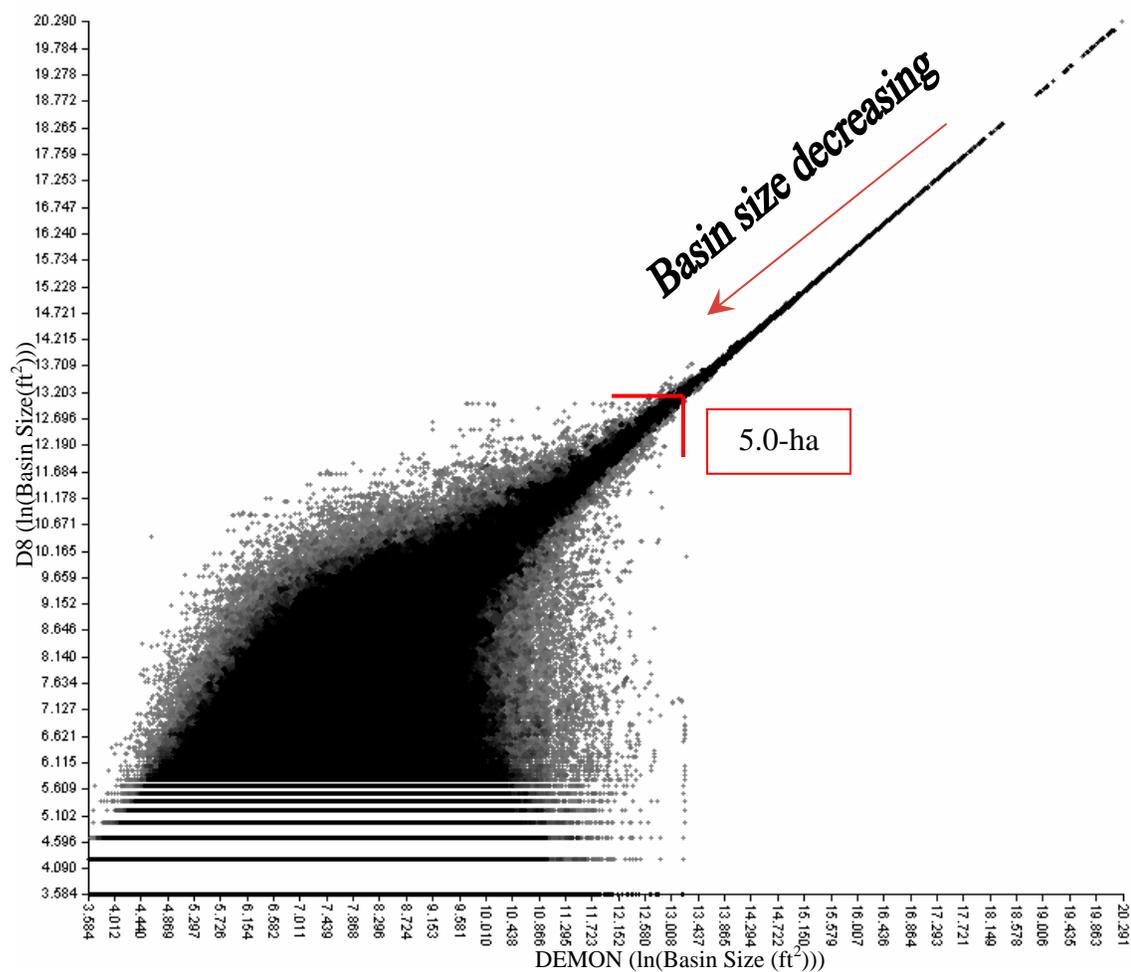


Figure 8. Cell plot of entire catchment area for a 2-m LiDAR DEM at the study site (Natural Log Values). D8 and DEMON correlate well above a threshold of approximately 5-ha as shown with the red mark. Below the 5-ha mark, differences are shown by the scatter.

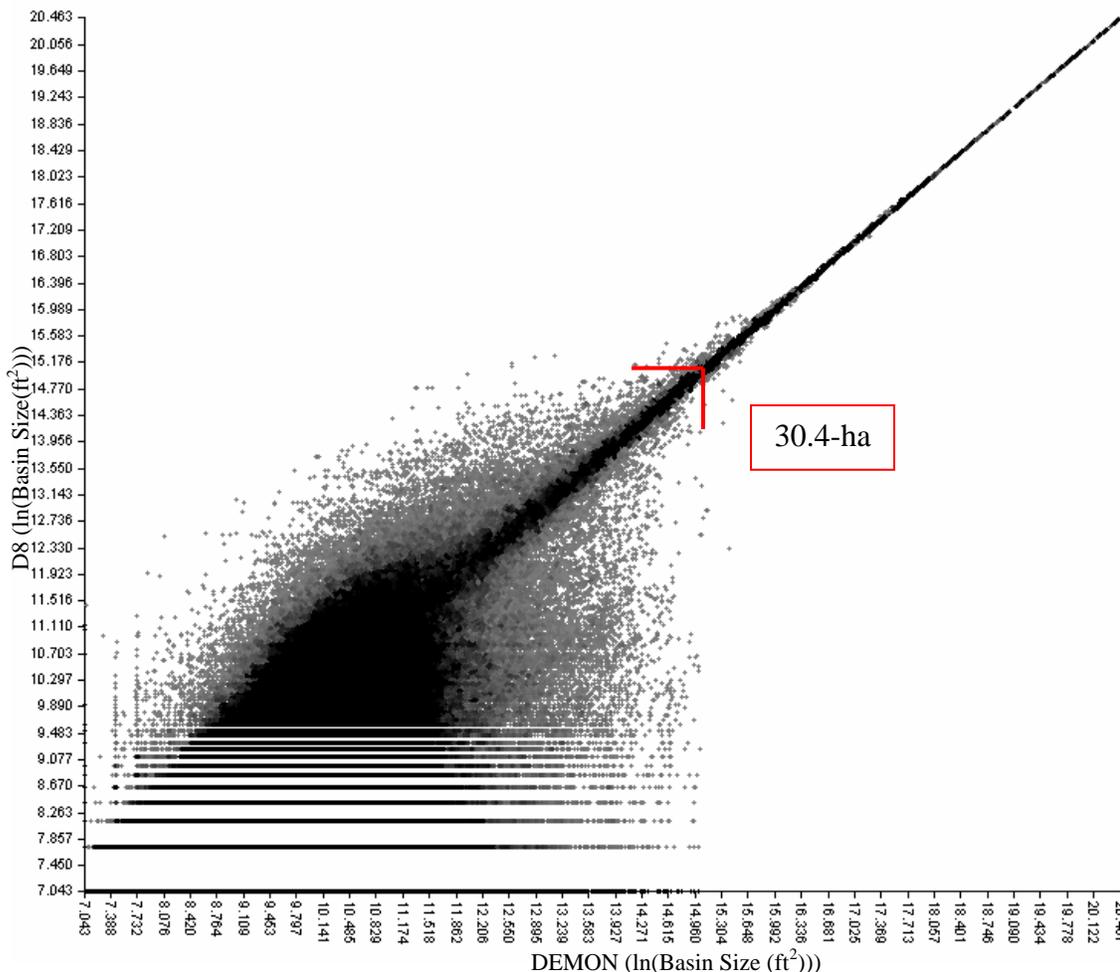


Figure 9. Cell plot of entire catchment area for the 10-m USGS at the study site (Natural Log Values). D8 and DEMON do correlate although not until a catchment area of approximately 30-ha is acceded as shown with the red mark. Below the 30-ha mark, differences are shown by the scatter.

Table 5. Relation between D8 and DEMON algorithms at various resolutions. As resolution is decreased, the correlation between D8 and DEMON becomes less. Basin size values determined by visually inspecting the plot of catchment area figures.

DEM Resolution	Basin Size (ha)
2-m LiDAR DEM	5.0
6-m LiDAR DEM	6.8
10-m LiDAR DEM	9.1
10-m USGS DEM	30.4

Using LiDAR datasets, D8 determines stream networks as well as DEMON. Endreny and Wood (2003) gave 2D-Lea (a building block in DEMON) the highest ranking in

accuracy in comparison to any stream model that was used in their study. The data suggested that increased DEM resolution decreased the need for sophisticated models, reducing processing times required by complex models for high-resolution DEM's. Since D8 is the most commonly used model and simplest to implement, computational time in computing stream networks is reduced in comparison to DEMON.

When comparing a LiDAR derived 10-m DEM with a USGS 10-m DEM, D8 stream channels with a catchment size of about 12-ha and greater somewhat converge between the two DEM's (Figure 10). When catchments are less than 12-ha, no convergence existed. Since a USGS 10-m DEM contained topographic errors in regard to stream channel location, streams from the LiDAR and USGS were categorized as identical if they were less than 90-m apart to decrease error between the two datasets. This caused differences to decrease significantly between the two stream networks (Appendix H). These differences occur in Strahler order 1,2 and 3 streams. The area in the lower middle of Figure 10 corresponds with areas when that the 10-m USGS under predicted stream channels in comparison to the area to the left side of the figure which corresponds to areas of under prediction of the LiDAR 10-m.

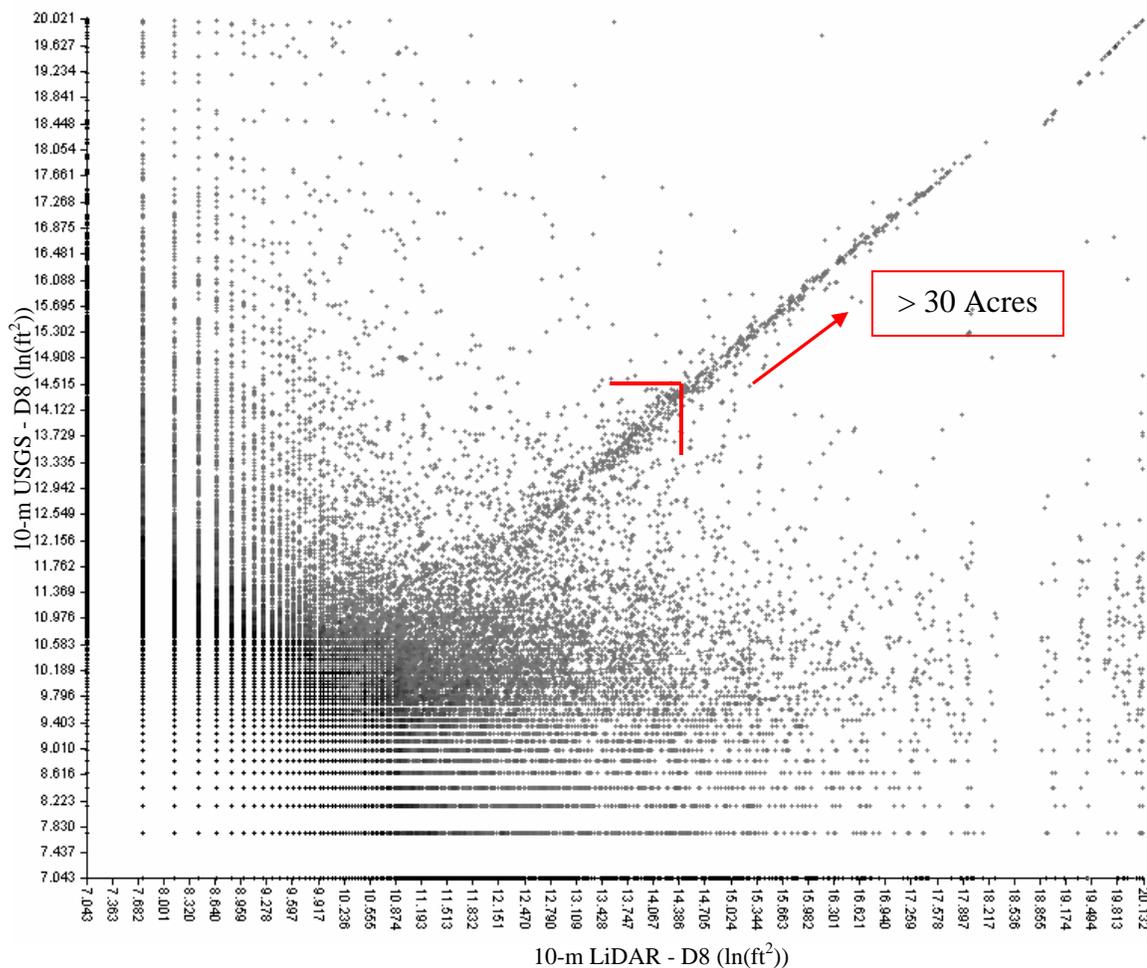


Figure 10. The D8 flow algorithm applied to the USGS and LiDAR Generated 10-m DEM. Illustrates that stream channels with a catchment size of about 12-ha and greater somewhat converge between the two DEM's with regards to D8. When catchments are less than 12-ha, differences in stream channel location are shown by the scatter.

3.2 Resolution Effects on Flow Direction

As the DEM resolution increased, D8 model sensitivity also increased. At the 2-m resolution, a road crossing a stream is seen as a dam therefore routing the stream onto the road (Figure 11). The road influence alters stream location and extent. To correct this problem, known culvert locations, stream or ditch culverts at the study site were used to make the stream continue under the road. The LiDAR DEM was lowered in elevation at the culvert locations to cause the stream channels to flow to the culverts and away from the road. Culvert data causes the streams to flow to the main channel thereby minimizing road effects (Figure 12) (Schiess and Tyrall, 2003).

Decreasing the LiDAR-DEM resolution to 6-m removed the road effect and placed streams in a more realistic location than the 2-m uncorrected. At 6-m resolution, the stream models could not identify roads or the ditches associated with the roads. As the LiDAR-DEM resolution decreased, road influence decreased. Stream channels, for the most part, followed the corrected 2-m stream network (Figure 13). The advantage of the 6-m LiDAR-DEM was that it provided a significantly improved stream network compared to the 10-m USGS DEM and removed the need for culvert data.

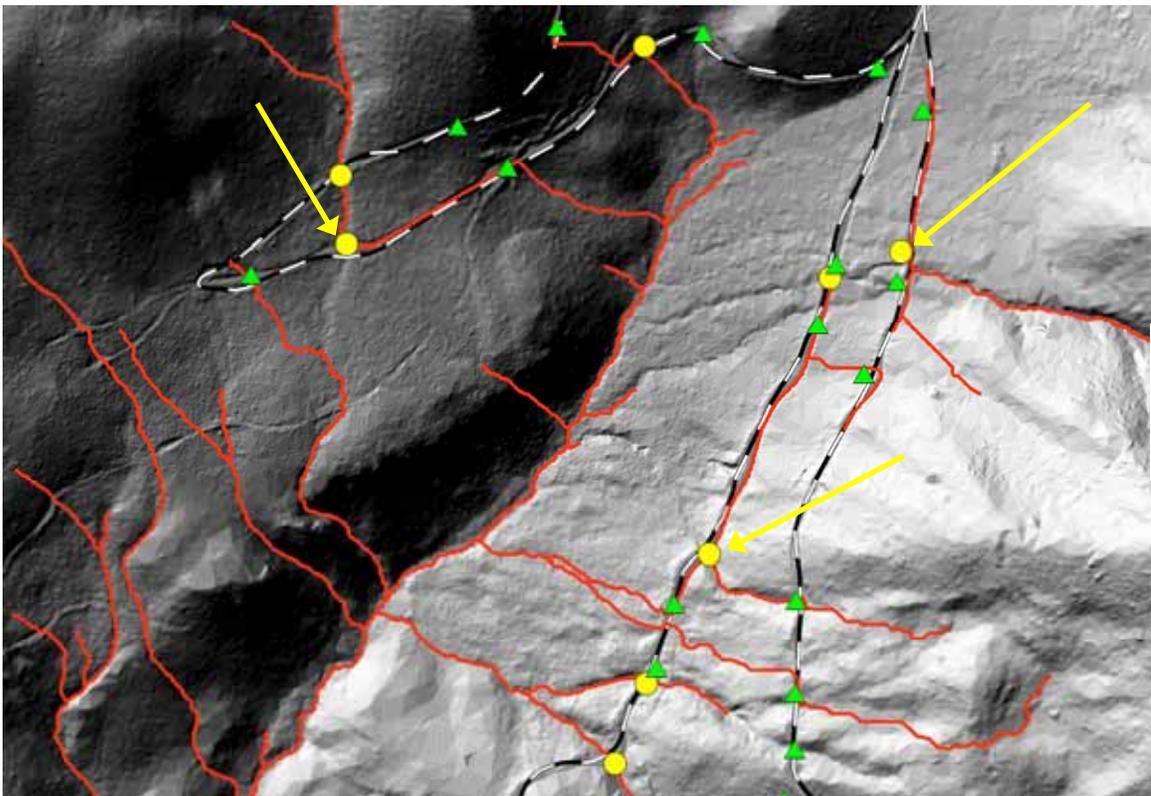


Figure 11. Streams generated from the 2-m LiDAR-DEM, in red, without using culvert correction. Stream culverts are circled, ditch culverts are triangles. At the 2-m resolution, the models defined some roads as stream channels by bypassing the stream culverts (arrows).

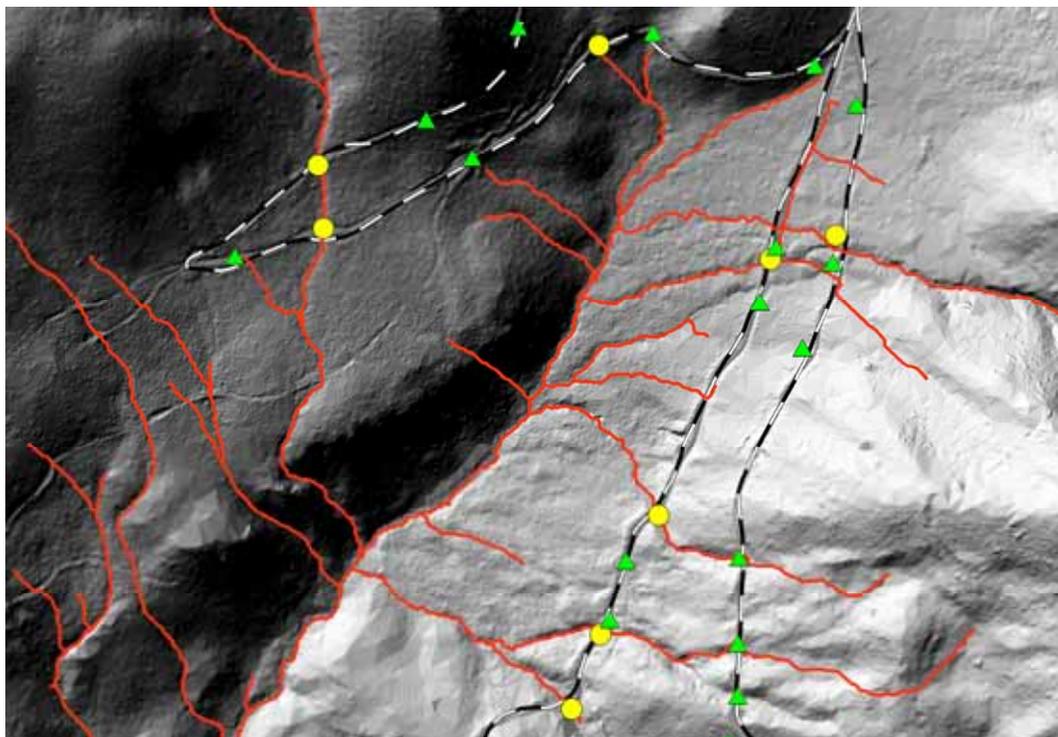


Figure 12. Streams generated from the 2-m LiDAR-DEM using culvert locations. Stream culverts are circled, ditch culverts are triangles. Culvert data causes the streams to flow downslope of the culvert allowing the stream to travel to the main channel more accurately.

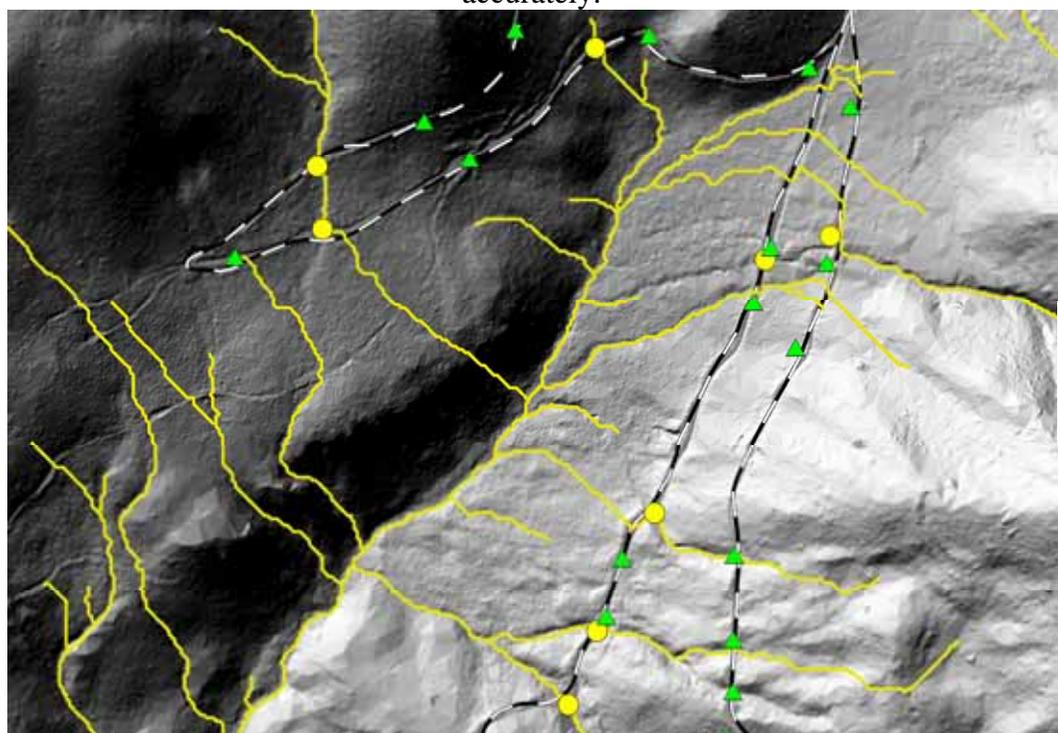


Figure 13. At 6-m resolution, the stream models did not route streams along roads and ditches. Removing the road effect placed streams in a reasonable location.

3.3 Assessing the Current Hydro Layer

Of the streams that the WA DNR identified as Type 9 (section 1.3), 72% were observed as not containing water. Half of those did not even contain an identifiable stream channel. The other half could be considered seasonal even though water erosion was not present. The remaining 28% that contained water also contained a perennial initiation point.

Very few of the streams typed as 5 were dry and most were perennial. Figure 14 illustrates field verified perennial streams identified using 6-m LiDAR DEM versus what the DNR considers perennial. Not all streams were visited due to stream head inaccessibility. There is a 530% increase in perennial stream length going from the DNR hydro layer perennial steams to the field verified perennial streams (Table 6). If a uniform buffer of 30 meters, a standard DNR regulation estimate, was placed around the stream channels, there would be a 350% increase in buffered area.

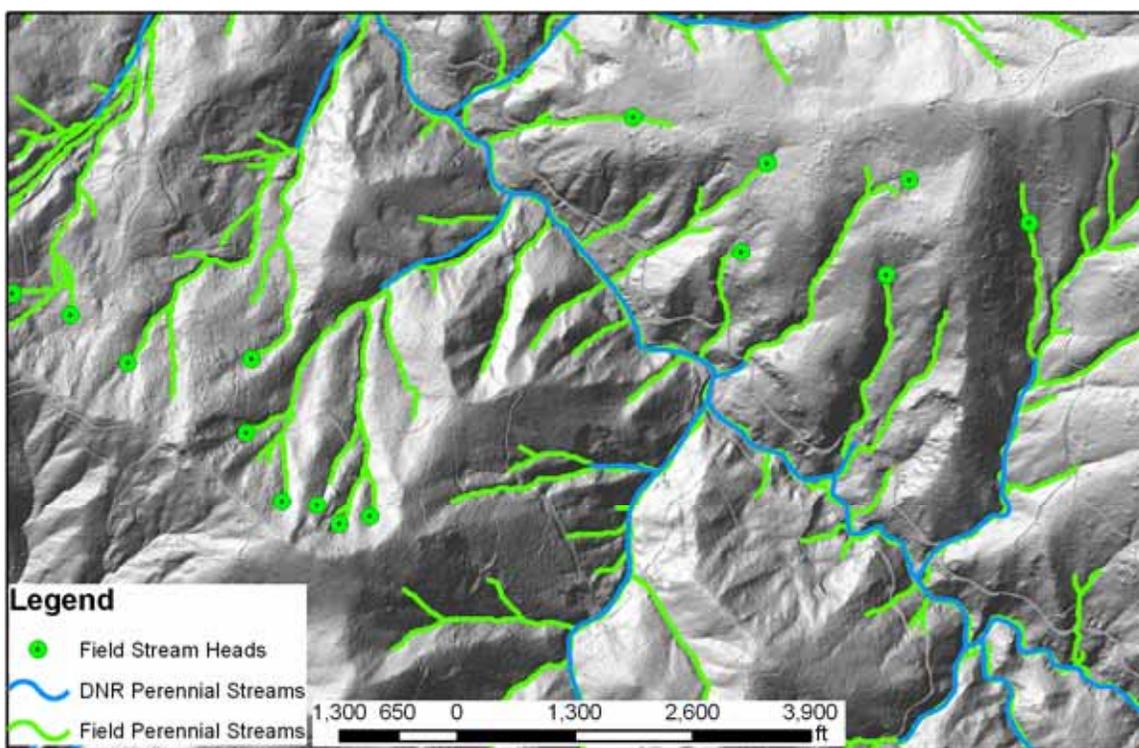


Figure 14. Field verified perennial streams using LiDAR in green vs. what the DNR considers perennial in blue.

Table 6. Differences in perennial stream length between DNR hydro layer and the LiDAR stream network derived from 6-m LiDAR DEM.

Perennial Stream Datasets	Length (km)	Buffer Area (ha) *
DNR Hydro Layer created using USGS 7.5'	68	240
LiDAR Streams	362	860

*Uniform 30-m buffer for both datasets for Perennial flow

3.4 Determining Perennial Streams

Given the soils geology and topography of the Tahoma State Forest, perennial flow began when ground water surfaced to form a stream head (Figure 15). Almost 90% of the stream heads located in the field fit this description interpreted as a spring (Figure 16).

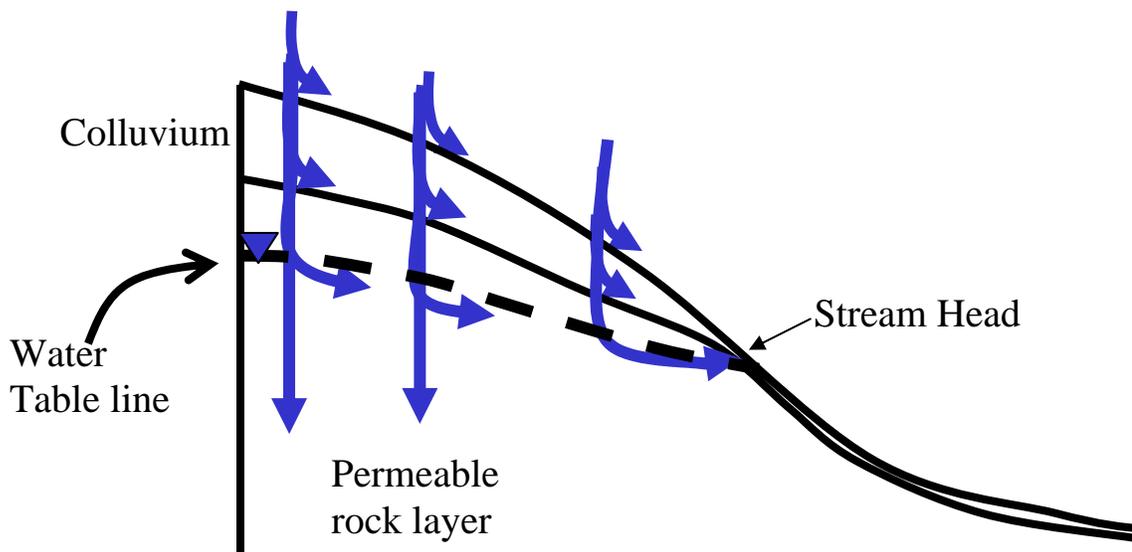


Figure 15. Stream head defined by the landscape. Perennial flow begins when ground water surfaced to form a stream head



Figure 16. Spring identified as a stream head in the field.

In the field, 61 stream heads were located within the Mineral Creek, North Fork watershed. Table 7 lists the distribution on the stream heads and Figure 17 displays the locations of the heads. Using the method described in the PIP Model section, 53 heads were selected at random within the study site to be used to create a model to predict perennial initiation points (PIP).

Table 7. Sub-Basins used in perennial head identification with the number of stream heads visited in the field.

Sub-Basin	Sub-Basin Area (ha)	Stream Heads Visited
1	318	12
2	473	13
3	366	23
4	486	4
5	526	5

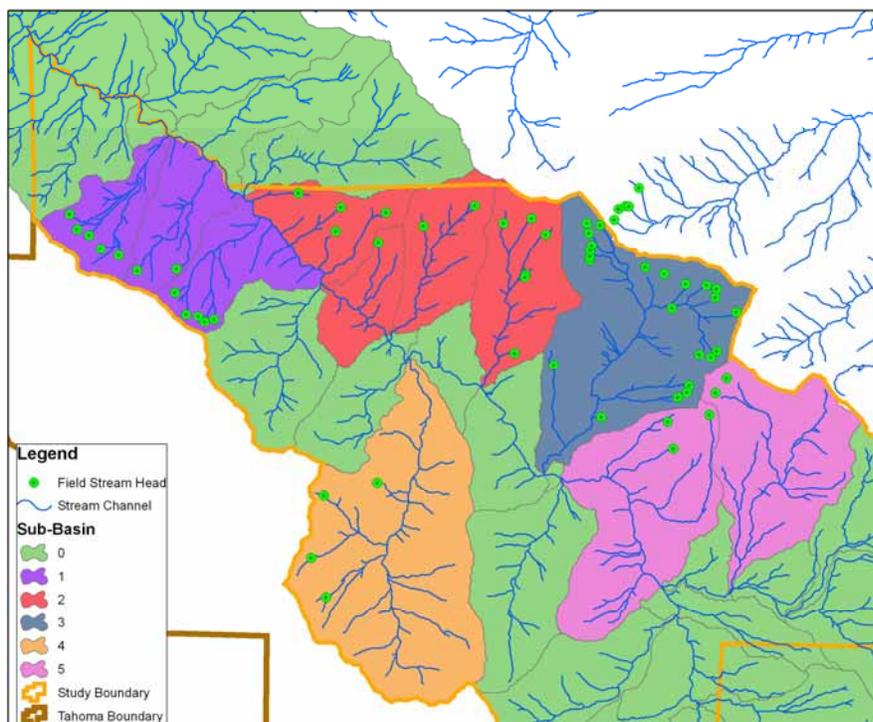


Figure 17. Stream heads in green within the study site. Stream channel generated from the 6-m LiDAR-DEM.

The final Linear Regression model for PIP used fewer variables than expected. The final model selected Basin Size using D8, Percent Slope, and Precipitation. Downstream gradient, forest density, elevation, and site class could not be used to create the equation for determining the probability of stream head locations based on a 0.05 significance level. Table 8 summarizes the variables used for the regression. The Hosmer-Lemeshow chi-square statistic for this model was 10.262 and the -2 Log likelihood statistic was 80.130. Self-classification accuracies for this model were 77.4% for perennial flow and 88.7% for non-perennial flow. APPENDIX I provides further statistics regarding regression.

Table 8. Summary of the final logistic regression model.

Coefficients	Estimate	Standard Error	Significance	Exp(B)	95.0% C.I. for Exp(B)	
					Lower	Upper
Log10(Basin Size)	7.235	1.425	0.000	1386.737	84.879	22656.314
Precipitation	0.477	0.184	0.010	1.612	1.123	2.313
% Slope	0.096	0.040	0.016	1.101	1.018	1.191
Constant	-45.172	16.050	0.005	0.000		

Using a model based on basin size alone for predicting perennial stream channels would be less accurate than the above model. The Hosmer-Lemeshow chi-square statistic for basin size model was 8.864 and the -2 Log likelihood statistic was 91.840. Self-classification accuracies for the basin size model were 77.4% for perennial flow and 84.9% for non-perennial flow. Overall, the average basin size for perennial flow for this model is 1.28-ha (3.16 acres).

Both models can over estimate the extent of perennial stream channels by placing flow upstream of the PIP. Using the conservative approach described in the PIP Model section it has the potential to under predict perennial flow. Using average basin size determined from the field data, the average value is 2.2-ha (5.44 acres). This average over and under predicts perennial flow. Since the Washington State Register defines contributing basin area as at least 21-ha (52 acres), all models and approaches would significantly correct WAC estimations.

The model in Table 8 was run at 4 different resolutions, 2-m, 6-m, 10-m LiDAR and a 10-m USGS DEM. Figure 18 illustrates the change in distance between modeled stream head location and field-verified stream head location for different DEM resolutions. This indicates that at all LiDAR resolutions, the error is relatively the same. The 10-m USGS DEM average distance and spread are higher than the LiDAR. This confirms that LiDAR improves upon modeling stream heads more accurately than a 10-m USGS DEM. The reasoning for high distances in the figure is due to the model not predicting a stream where the field verified stream head was located.

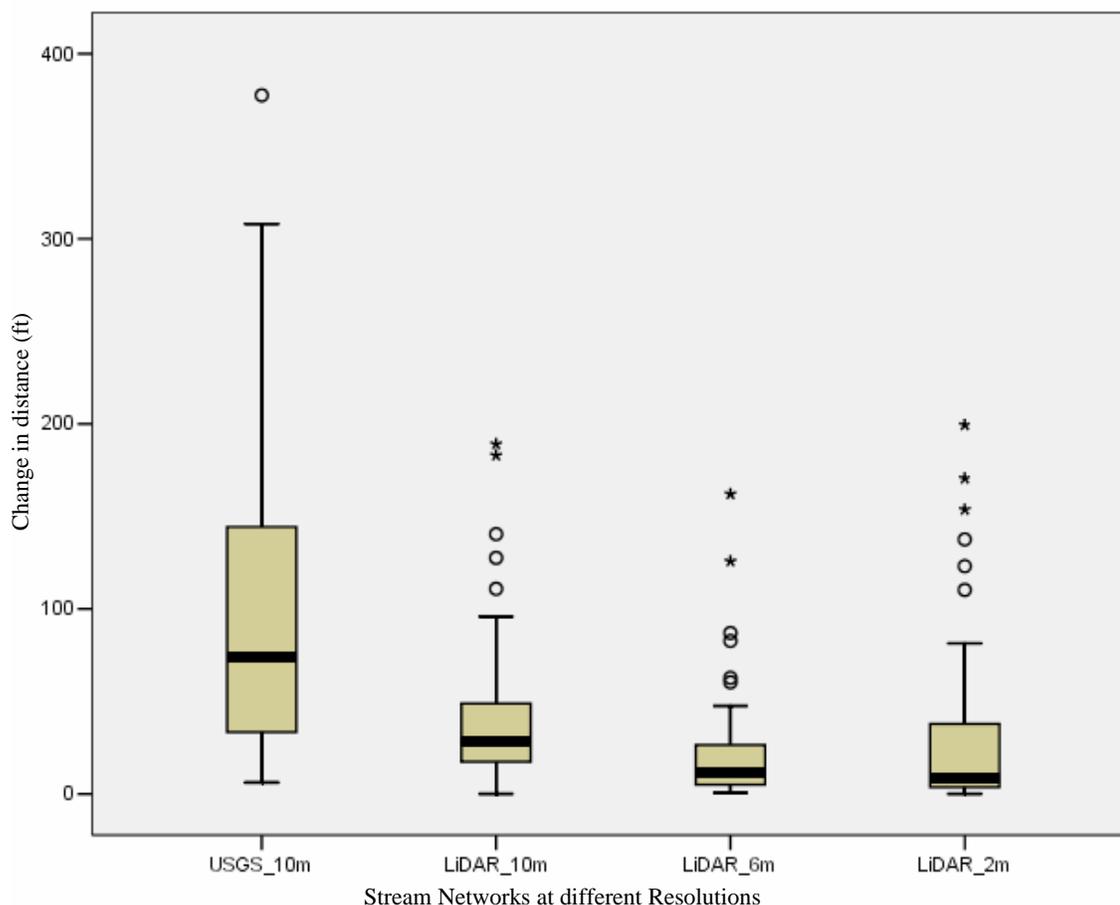


Figure 18. The distance error between modeled stream head location and field-verified stream head location at a given resolution. This indicates that at all LiDAR resolutions, the error is relatively the same. The 10-m USGS DEM average distance and spread are significantly higher than the LiDAR.

The PIP Model was tested on the various flow direction techniques listed in the “Flow Direction Methods Utilized” section to test which flow direction algorithm worked best in locating perennial flow. Because of the differences in the MFD and Dinf from D8, a separate bilinear regression model was created but none of the variables were significant based on a 0.05 significance level. The regression model in Table 8 was then used on the various algorithms on a 6-m LiDAR DEM to see the errors in predicted stream head locations. DEMON and D8 stream heads correlated in the difference in distance from the field-verified stream heads. Dinf and MFD increased in the difference in distance from field stream heads when compared to D8 (Figure 19). Finding a way to develop a

bilinear regression model for Dinf and MFD would reduce the data error illustrated below.

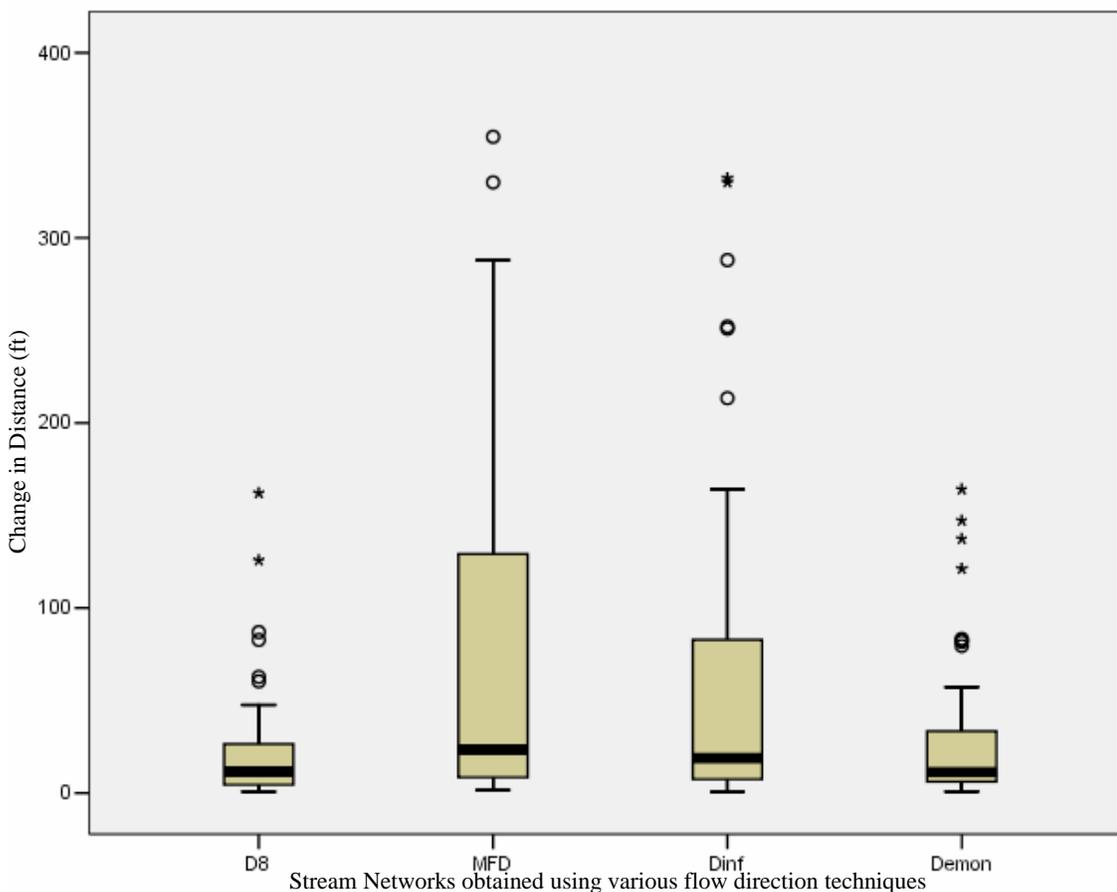


Figure 19. The distance error between field-verified stream heads and various flow direction modeled streams. DEMON and D8 correlated in error while Dinf and MFD significantly increased in error.

3.5 Determining Fish Stream Habitat

The CMER Model and Gradient Model estimated fish extent differently. CMER Model used basin area, downstream gradient, elevation, and precipitation, while the Gradient Model uses only downstream gradient. With field verification, the Gradient Model located fish barriers providing accurate fish habitat maps. The CMER Model tended to place potential fish waters well upstream of waterfalls and culvert barriers. Overall, the Gradient Model predicted fish extent closer to the main channel than the CMER method.

Figure 20 displays a longitudinal profile of a selected creek generated by a LiDAR DEM. The red zones indicate field verified waterfalls and culvert/road locations that the Gradient Model identified. These waterfalls can range from 1 to 6+ meters tall. If trout were able to pass these barriers, the CMER Model would be correct in its estimation. Figure 21 shows the predicted fish habitat estimated by the Gradient Model. The CMER Model and DNR Hydro layer are within the predicted fish habitat area but overlook small, but critical segments that act as fish barriers. Gaps in between the green zones are the waterfall and culvert barriers. These results were similar to the results for other stream channels throughout the study site.

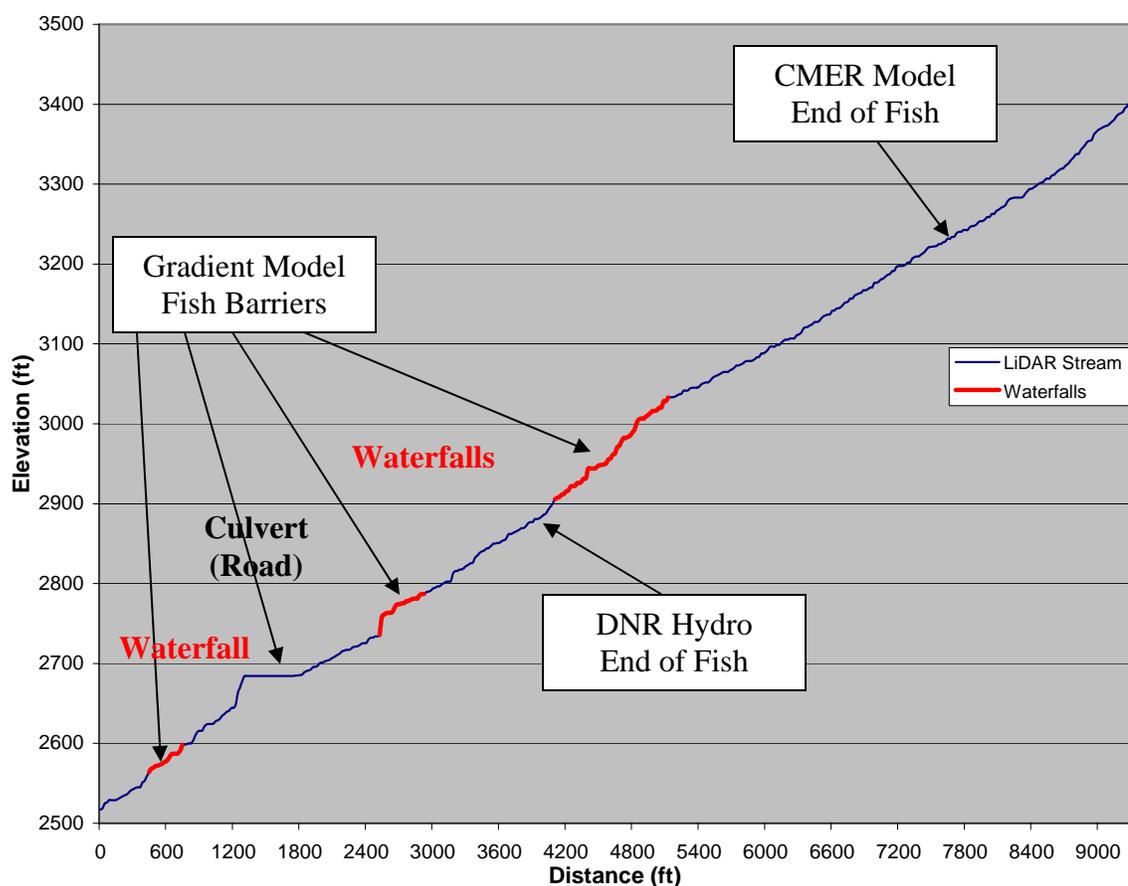


Figure 20. Longitudinal Profile of a selected creek. (Distance from main channel) The red zones indicate field verified waterfalls that the Gradient Model identified as well as culvert/road locations.

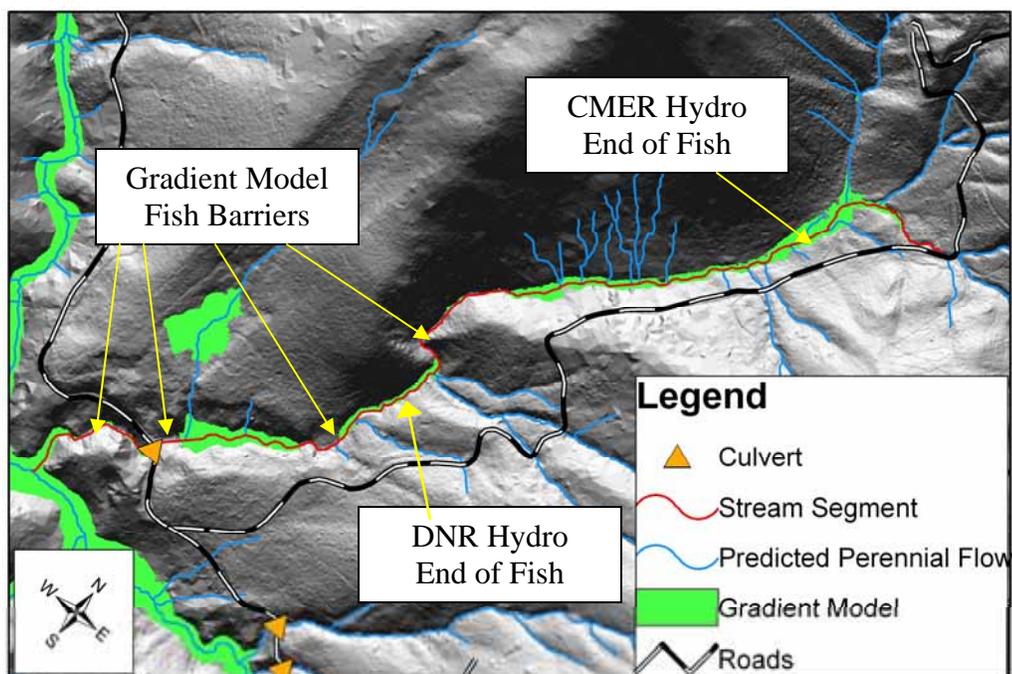


Figure 21. Stream Channel used (red) from Figure 20. This shows the predicted fish habitat in green estimated by the Gradient Model. The CMER Model and DNR Hydro layer are within the predicted fish habitat area but overlook small, but critical segments that act as fish barriers. Gaps in-between the green zones are the barriers.

The CMER Model and the DNR Hydro layer predicted fish presents further up stream than the Gradient Model. Using the Gradient Model, fish-bearing streams decrease 24% from the DNR Hydro layer (Table 9). The CMER Model potentially overestimates fish-bearing streams by 54% when compared to the Gradient Model.

Table 9. Predicted Fish-Bearing Streams within the Study Site using different techniques.

Datasets for Fish-Bearing	Length (km)
DNR Hydro Layer	18.0
CMER EOFP Model	22.4
Gradient Model using LiDAR Streams	14.5

The LiDAR DEM modeled fish barriers more accurately than a 10-m USGS DEM (Figure 20). At times, the USGS dataset resulted in incorrect locations for major waterfalls. In many cases, major waterfalls were not identified at all. For the most part, the USGS DEM identified the larger streams at the study site as accurately as the LiDAR

DEM. This makes sense considering how the USGS DEM was created. Large streams are visible from aerial photos allowing elevation readings to be fairly accurate.

3.6 Weaknesses and Shortcomings

Having a LiDAR DEM dataset provides the ability to model stream channels with their respective stream head location more accurately compared to using the standard 10-m USGS DEM. LiDAR DEMs also allow for more accurately locating potential fish barriers. Currently, less than 1/5 of Washington State has LiDAR coverage limiting the ability to model stream networks across large portions of the state. The DNR as well as the timber industry appear to be set on rapidly expanding their LiDAR coverage which would eliminate this problem.

Current spatial information usually does not have the fine resolution of a LiDAR dataset. Soils surveys, geology, and site class information are at lower resolutions than LiDAR so trying to develop a model that uses those themes could be difficult. These datasets would need to be at a higher resolution to better match a LiDAR derived dataset. The lack of a high-resolution site class map may explain why site class could not be used in the perennial stream regression model.

The major limitation to the regression model used to predict perennial flow is that it is site specific. While it proved to be more accurate than the current datasets available, accuracy may decrease when applying the model to a variety of landscape types. Data collected on a larger, more diverse area would generate a model better able to model landscapes.

In the middle of Figure 21 just below the “CMER Hydro End of Fish” label, a series of stream channels were identified in blue. This is a misrepresentation of the landscape due to dense forest canopy. A dense canopy will limit the amount of laser ground returns causing a loss of resolution in specific locations. A way to correct this problem would be to conduct field surveys of the landscape at those locations.

4. CONCLUSION

When used with high-resolution DEMs, the most common and least sophisticated algorithm to locate stream channels based on flow, D8, proved to be sufficient when compared to more complicated and process demanding algorithms. The advantage of D8 is that most programs like ArcGIS come standard with this algorithm to locate stream networks. In generating stream data using LiDAR, a computer with a high-speed processor is not necessary.

Increasing surface topographic resolution allowed an increase in the precision of stream channel modeling. Since the LiDAR DEM that was used for this project was at a 2-m resolution, decreasing the resolution of the LiDAR did not alter the ability to model locations of stream channels. The resolution at 2-m proved sensitive for the models used requiring a culvert dataset to alter stream flow off the road networks. Decreasing resolution to 6-m eliminates road effect errors. Relying on the resolution of a 10-m USGS DEM proved inadequate in modeling head water streams. High-resolution DEMs are necessary to accurately model stream networks.

As shown, a LiDAR DEM is a great source for generating hydrologic data by identifying probable perennial stream channels and locating fish barriers along a given stream. The ability to accurately model headwater streams and identify fish-bearing streams allows stream buffer zones to be in their topographically correct location. In turn it will allow for better planning at both the strategic (sustainable harvest volume calculations) and tactical or map scale level (forest operations planning at the watershed level) rather than having to rely on time-consuming field reconnaissance.

REFERENCES

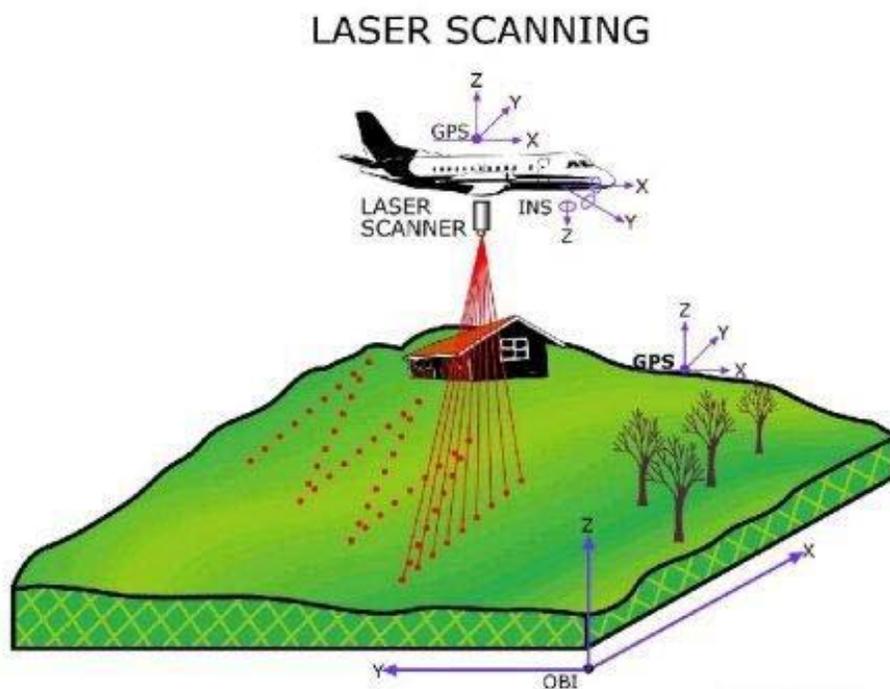
- Conrad, R.H., B. Fransen, S. Duke, M. Liermann, and S. Needham. 2003. The development and assessment of the preliminary model for identifying fish habitat in western Washington. Draft report prepared for Washington Dept. Natural Resources Instream Scientific Advisory Group.
- Costa-Cabral, M. C. and S. J. Burges. 1994. Digital elevation model networks (DEMON): A model of flow over hillslopes for computation of contributing and dispersal areas. *Water Resources Research*, 30, 6, 1681-1692
- Daly, C., C. Taylor, and G. Taylor. 1988. 1961-1990 Mean Monthly Precipitation Maps for the Conterminous United States. Oregon State University Spatial Climate Analysis Service
- Daly, C., C. Taylor, and G. Taylor. 2000. 1971-2000 Mean Monthly Precipitation Maps for the Conterminous United States. Oregon State University Spatial Climate Analysis Service
- Damian, F. 2001. Improving cross drain culvert spacing with GIS interactive design tool. The International Mountain Logging and 11th Pacific Northwest Skyline Symposium.
- Endreny, T. A., and Wood, E. F. 2003. Maximizing spatial congruence of observed and DEM-delineated overland flow networks. *Int. J. Geographical Information Science*. Vol. 17, No. 7, February 2003, pp. 699-713
- Endreny, T. A., and Wood, E. F. 2001. Representing elevation uncertainty in runoff modeling and flowpath mapping. *Hydrologic Processes*. 15, 2223-2236 (2001)
- Haugerud, R.A., and D. J. Harding. 2001. Some algorithms for virtual deforestation (VDF) of LiDAR topographic survey data. *International Archives of Photogrammetry and Remote Sensing*. XXXIV-3/W4, p. 211-217.
duff.geology.washington.edu/data/raster/LiDAR/vdf4.pdf (March 2003).
- Jenson S. K. and J. O. Domingue. 1988. Extracting topographic structure from digital elevation data for geographic information systems analysis. *Photogrammetric Engineering and Remote Sensing*. Vol. 54, No. 11, pp. 1593-1600
- Latterell, J. J., R. J. Naiman, B. R. Fransen, and P. A. Bisson. 2003. Physical constraints on trout (*Oncorhynchus* spp.) distribution in the Cascade Mountains: A comparison of logged and unlogged streams. *Canadian Journal of Aquatic and Fisheries Sciences* 60:1007-1017.
- Luijten, J, 2000. Dynamic hydrological modeling using ArcView GIS. *ArcUser*, July-Sept. 2000. <http://www.esri.com/news/arcuser/0700/hydro.html>

- Montgomery, D. R. and W. E. Dietrich. 1992. Channel initiation and the problem of landscape scale. *Science*, 255: 826-830.
- Montgomery, D.R., and E. Foufoula-Georgiou. 1993. Channel network source representation using digital elevation models. *Water Resources Research*, 29, 12, 3925-3934
- Moore, I. D., R. B. Grayson and A. R. Ladson. 1991. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5(1): 3-30.
- Moore, I. D. and R. B. Grayson. 1991. Terrain-based catchment partitioning and runoff prediction using vector elevation data. *Water Resources Research*, 27, 6, 1177-1191
- North Carolina Flood Plain Mapping Program. 2003. LiDAR and digital elevation data. Jan. 2003, http://www.ncfloodmaps.com/pubdocs/LiDAR_final_jan03.pdf. Accessed January 10, 2005
- O'Callaghan, J. F., and D. M. Mark. 1984. The extraction of drainage networks from digital elevation data. *Comput. Vision Graphics Image Processing*, 28, 323-344
- Oladi, D. 2001. Developing a forest growth monitoring model using thematic mappery imagery. *Proc. ACRS 2001 - 22nd Asian Conference on Remote Sensing*, 5-9 November 2001, Singapore. Vol. 2, pp. 1457-1462
- Pan, F., C. D. Peters-Lidard, M. J. Sale, A. W. King. 2004. A comparison of geographical information systems-based algorithms for computing the TOPMODEL topographic index. *Water Resource Research*, 40.
- Quinn, P. F., K. J. Beven, P. Chevallier, and O. Planchon. 1991. The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. *Hydrol. Processes*, 5, 59-80
- Renslow, M. S. 2001. Development of a bare ground DEM and canopy layer in NW forestlands using high performance LiDAR. *Int'l User Conf., ESRI*.
- Schiess, P. and A. Mouton, 2005. North Fork Mineral Creek access and transportation strategy. *Techn. Report, University of Washington, WA. Spring 2005*
- Schiess, P. and J. Tyrall, 2002. DNR Hydro and LiDAR streams areas in buffers, Hydro and LiDAR
http://courses.washington.edu/fe451/projects/02_tyee/Presentation/Presentation_2002_files/frame.htm

- Schiess, P. and J. Tryrall, 2003. Developing a road system strategy for the Tahoma State Forest. Techn. Report, University of Washington, WA.
- Schiess, P. and Krogstad, F. 2003. LiDAR-based topographic maps improve agreement between office-designed and field-verified road locations. Proceedings of the 26th Annual Meeting of the Council on Forest Engineering,, Bar Harbor, Maine, USA, 7-10 September, 2003
- Shamayleh, H., and A. Khattak. 2003. Utilization of LiDAR technology for highway inventory. Mid-Continent Transportation Research Symposium, Ames, Iowa, August 21-22, 2003.
- Soil Survey Staff. 1998. Official series description – Bellicum, Cattcreek, and Cotteral series. Natural Resources Conservation Service, United States Department of Agriculture. Retrieved May 6, 2005 from <http://ortho.ftw.nrcs.usda.gov/cgi-bin/osd/osdname.cgi>
- Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. Transactions of the American Geophysical Union 38, 913-920.
- Tarboton, D. G., R. L. Bras, and I. Rodriquez-Iturbe. 1991. On the extraction of channel networks from digital elevation data. Hydrol. Proc., 5, 81-100
- Tarboton, D. G. 1997. A new method for the determination of flow direction and upslope areas in grid digital elevation models. Water Resour. Res., 33(2), 309-319
- Tarboton, D. G. 2003. Terrain analysis using digital elevation models in hydrology. 23rd ESRI International Users Conference, San Diego, California, July 7-11 2003
- Terrapin Environmental. 2003. Water typing model gield validation study design approach and procedures. Westside Water Type Model Validation Plan, Dec. 8, 2003
- Wilson J. P. and J. C. Gallant (editors). 2000. Terrain analysis: Principles and Applications. New York, John Wiley and Sons
- Wolock, D. M., and G. J. McCabe. 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters. Water Resour. Res., 31(5), 1315-1324

APPENDIX A: LiDAR System Used

LiDAR (Light Detection And Ranging) data is collected by a well defined flight plan through a specified location. (Appendix Figure A1) LiDAR is a method of detecting distant objects and determining their position or other characteristics by analysis of pulsed laser light reflected from their surfaces. LiDAR works on the same principle as RADAR (Radio Detection And Ranging), but LiDAR uses light waves emitted by a laser (rather than radio waves) to gather data. In its simplest form, LiDAR is used to determine the distance from the laser to a given object.



Appendix Figure A1. LIDAR Data Collection (Renslow 2001)

The data for the study site was collected by TerraPoint which flew a multiple-return scanning laser altimeter in a small fixed-wing aircraft with a circa 0.9 meter on-the-ground laser spot, nominal across- and along-track pulse spacing of 1.5 meters, and 50% overlap of adjacent flight lines, providing an average of circa 1 pulse/square meter. Some of these data (Bainbridge Island) were acquired for Kitsap PUD in 1996-1997. Average pulse spacing of the Bainbridge data was similar.

The data are in Stateplane projection, Washington South zone FIPS zone 4602. The vertical datum is NAVD88, horizontal datum is NAD83 HARN. Horizontal units are US Survey Feet. Raster cells (grid cells, image pixels) are 6 ft square. Elevations are recorded in integer feet.

APPENDIX B: LiDAR Data Collection

This section was developed as data was processed for the University of Washington Forest Engineering Capstone course using data provided by Washington State Department of Natural Resources. The LIDAR data that was being processed was for the Tahoma State Forest and was flown by TerraPoint, LLP in the spring of 2003.

It was decided to use an algorithm developed by Haugerud and Harding 2001 for a variety of reasons. The first was that the process is published and publicly available. Their connection with the University of Washington made access to assistance a little easier. The algorithm was also developed using data provided by TerraPoint, LLP so it had been used before and some of the issues related to processing on TerraPoint data were documented.

The data was provided in the Washington State Plane Nad83 as ASCII files. The data contained up to four returns per point. Before beginning the process of removing the vegetation points, Hans Anderson of Precision Forestry ran the data through a script that removed all of the points leaving only the last return for each point. The process described below is based on using the last return only data.

Process

1. The last return data files received from Hans were text files where the first three numbers were [X Y Z], space delimited. Using java application asc2gen.exe, the data was converted into Arc Info Generate format (see Appendix B). This creates new files in the directories with the same names as the input but with the extension of TXT_L.
2. Copy the file TXT2TIN.AML (see Appendix C) to the folder containing the text files of Generate and last return. Run this script. No arguments are needed.

3. A folder with the [name]_1 for each of the TXT_L files will appear. These are the last return TINs. Copy these folders to a new folder for Deforestation. The folder name GROUND was used and the name will be used for the rest of the process.
4. In this new folder ground, copy AML scripts VDF.AML, DESPIKE3.AML, ELAPSEDTIME.AML and TIN2DEM.AML (see Appendix D).
5. Change the working directory in Arc to the GROUND folder and run VDF.AML. The script will look for all the TINs in the directory and process them all at once. This is a very long process. See the following discussion for issues and things to watch out for when running the VDF scripts.
6. The final stage of the processing is to run the script TIN2DEM.AML. This is the script that handles the combining of the tins into one and then creating the grid and DEM files to be used in other applications. See the following discussion for issues and things to watch out for when running the TIN2DEM.AML script.

Discussion

The first time that this process was run, there were a few things that came up that are worthy of discussion. The first issue was time to process. It took about 12 days to process 270 files in 2 GB of ASCII data. Part of that time was used in modifying scripts to automate the processing of the files. Most of this time was in handling the creation and re-creation of all of the TIN files. Arc Info is very slow in processing these files. Due to a noticed problem in the DZ2 setting (discussed next) it was necessary to rerun the data. This time the process was run using scripts created by Hans Anderson in IDL. The processing time was dropped to about four days. Unfortunately IDL can only be run on a computer with the interpreter so it is not possible to use this standalone. It does leave some thoughts on ways to program the process into a standard language such as Java, C++, or Visual Basic in order to create the ground data set or at least reduce the data size that Arc Info needs to process.

The next problem that was noticed is that in areas that appeared to be dense stands and in particular younger stands, the scripts were not removing enough points so the areas were

showing up as raised areas. The problem was traced to the DZ2 setting in the VDF.AML file. The default setting from the website was 8' which is described as four times the grid size used in the processing. Since the default two-foot setting was left alone, it was thought this would stay the same. This setting was placed in the program to get rid of false minimum points in the data and would remove any points that were more than eight feet below the average for that grid. Unfortunately with very dense vegetation the majority of the point get caught high in the canopy and the few points that do make it to the ground are viewed as false lows. To fix this error the DZ2 setting has been modified to 100. This still allows for an initial check for extreme false lows while not removing points that are approximately tree height from the average.

One problem that was noticed that may not have a direct solution is areas of large triangle surfaces. These were particularly noticeable in the dense stands where a lot of the points never make it to the ground. This creates visible indications but has not been determined to provide any negative aspects to some of the analysis procedures like hydrology and view sheds. It does leave a reminder; in those areas the data is of lower quality at the ground level.

Managing the scripts and running them at the correct time was inconvenient at times. With just a little work it would be possible to create a master AML that handled all of the moving of data to the correct locations and running the correct AML to process the data. Due to the amount of data for Tahoma, it was broken into four blocks to be processed. This meant continually monitoring the processes to see where they were.

Most of the variables in the VDF aml were not changed. The one constant in DESPIKE3 that is worthy of a brief note is the two-foot grid setting in the Arc TINLATTICE command. The setting was left at two since the expected data quality was one point per square meter. The effect of increasing this value is not known but might be a way to speed up the processing by increasing the grid size. This would most likely decrease the ground quality to some degree but is something to look into.

A final caution when processing files in Arc Info, Arc is limited in the file name sizes. If any problems occur with the scripts stopping then the first thing to check is the names of the files. If the names are too large it will be necessary to rename the files. A simple script was created in DOS to automate the renaming process for all of the files that were attempted to manage.

File size is important when processing the entire LIDAR data using the TIN2DEM script. The processing limitation in Arc Info is 2.1 GB. Since the LIDAR set for the Tahoma State Forest combined was about 2.8 GB, this lead to Arc Info processing 75% of the data set. Arc Info would give an error message and leave the script during the creation of the MASTER_TIN.

Arc Info usually would allow the appending of the files prior to the creation of MASTER_TIN but at time would give an error messages indicating that it could not append the desired files.

The way to correct the problem was to run the TIN2DEM script for the north and south side of Tahoma State Forest separately. In order to merge the sides together, it is necessary to process an overlap section between the north and south. The outcome would be MASTER files for the north, south and middle section. Load the MASTER_GRD for the middles section into ArcMap and cut the edge around the grid out in order to remove the areas that were interpolated with only a few points. Use the merge function in raster calculator to merge the middle, north and south. Make sure to list the middle section first so that the boundary of the north and south will be ignored. Run the Hillshade process in ArcMap to better visualize if any error occurred in the merge. A visual error is a line from the boundary of the north, south, and middle. A minor error may not be avoidable.

APPENDIX C: GEN2ASC.EXE

The following bit of java code was written by to facilitate the processing of multiple ASCII (X, Y, Z) files into Arc Info Generate format. The program creates a small interface that allows window control of selecting files. In any given folder select as many files as desired of the correct format and the files will be formatted automatically and the new files will be placed in the same directory with a "txt_1" extension.

```

/*
 * @(#)asc2gen.java
 * Matthew Walsh
 * March 12, 2003
 *
 * Reads files in format "x, y, z" and creates an Arc Generate file
 *
 */

import java.awt.*;
import java.awt.event.*;
import javax.swing.*;

class asc2gen extends JFrame {
    private JPanel buttonPanel;
    private datman data;
    private JButton function3;

    public asc2gen() {
        super("Ascii(x,y,z) to Generate");
        data = new datman(); // instantiate the data controler

        function3 = new JButton( "Convert Files" );
        function3.addActionListener(
            new ActionListener() {
                public void actionPerformed((ActionEvent e)
                {
                    data.openFile();
                }
            )
        );

        buttonPanel = new JPanel();

```

```

        buttonPanel.setLayout( new GridLayout(2,4) );
        buttonPanel.add( function3);

        Container c = getContentPane();
        c.add( buttonPanel, BorderLayout.SOUTH );

        setSize( 580, 700);
        show();
    }

    public static void main(String args[]) {
        asc2gen app = new asc2gen();

        app.addWindowListener(
            new WindowAdapter() {
                public void windowClosing(WindowEvent e)
                {
                    System.exit(0);
                }
            }
        );
    }
}

/*
 * @author    Matthew Walsh
 * @(#)datman.java    03/04/03
 *
 * Handles the reading of X,Y,Z format ascii data and writing ArcInfo Generate files.
 */

import java.io.*;
import java.util.*;
import javax.swing.*;

class datman {
    public String x, y, z;
    public int numCommands=0, curCommand=0;
    private BufferedReader input;
    private BufferedWriter output;
    private JFileChooser fd;

    public datman() {

        fd = new JFileChooser();
    }
}

```

```

        fd.setSelectionMode( JFileChooser.FILES_ONLY );
        fd.setMultiSelectionEnabled(true);
    }

/**
 * Calls JFileChooser and handles which methods to call in order to
 * read the files.
 */
public void openFile( )
{
    String stRead = ""; // Line strings for running StringTokenizer
    String stCurToken = ""; // string containing current token for parsing.
    String theCommand = "";

    int result = fd.showOpenDialog( null );

    // check for cancel button
    if ( result == JFileChooser.CANCEL_OPTION ){
        return;
    }

    File fileList[] = fd.getSelectedFiles();

    for (int k = 0; k < fileList.length; k++) {
        File fileName = (File) fileList[k];

        if (fileName == null || fileName.getName().equals("")){
            JOptionPane.showMessageDialog( null, "Invalid File Name",
"Invalid File
name", JOptionPane.ERROR_MESSAGE );
        }
        else {
            // open the file
            try {
                if (numCommands != 0)
                    this.closeFile();

                input = new BufferedReader(new FileReader (fileName),
10240);
                //
                JOptionPane.showMessageDialog( null,
fileName.getName(),
"File Name", JOptionPane.INFORMATION_MESSAGE );

                output = new BufferedWriter(new
FileWriter(fileName.getName()+ "_1"), 10485760);

```

```

for (int i = 0; i < fileName.length(); i++) {

    stRead = input.readLine();
    StringTokenizer st = new StringTokenizer(stRead,
" ");

    x = st.nextToken();
    y = st.nextToken();
    z = st.nextToken();
    // System.out.println("x: " + x + " y: " + y + " z: " +
// z);

    output.write("1 " + x + " " + y + " " + z);
    output.newLine();
    // output.flush();

}

    output.write("END");
    output.flush();
    output.close();

}
catch (IOException e ) {
    JOptionPane.showMessageDialog( null, "Error Opening
File",
"Error", JOptionPane.ERROR_MESSAGE );

}
catch (NullPointerException e) {
    // placed here just to catch the times when the program
attempts to read too far.
    System.out.println("Null Point error in data reading");
    try {

        output.write("END");
        output.flush();
        output.close();
    } catch (IOException f) {
        System.out.println("Null Point error on the second
try/catch");
    }

}

}

}

}

```

```
/**
 * Closes the data files and removes the data from memory.
 */
public void closeFile( )
{
    try {
        input.close();

        this.numCommands = 0;
        this.curCommand = 0;

    }
    catch (IOException ex){
        JOptionPane.showMessageDialog(null, "Error closing file",
"Error",
JOptionPane.ERROR_MESSAGE );
        System.exit(1);
    }
}
}
```

APPENDIX D: TXT2TIN.AML

This is the file that creates the first TIN, triangulated surface from the point data, which will be used in the rest of the Virtual deforestation process. Run this aml through Arc in the directory that has the last return txt and txt_1 files. The script automatically selects all of the files and converts them into tins with the [name]_1 and no extension. When viewed in the windows browser they will show up as folders.

```

/* Following segment was added by Matthew Walsh
/* 13 March 2003
/* Allowed the automation of many tin files in one directory
/* Update of original files from Ralph Haugerud of USGS

&sv file [filelist *.txt last.list -file]

/* opens file list
&sv fileunit [open last.list openstat -read]
&if %openstat% ne 0 &then
    &return Problem opening file

/* reads file list
&sv tilefull [read %fileunit% readstat]
&if %readstat% ne 0 &then
    &return Problem reading file

&do &while %readstat% = 0
    &sv tile [entryname %tilefull% -noext]
        &if [exists %tile%_1 -tin] &then kill %tile%_1 all
        createtin %tile%_1 0.01 0.01
        generate %tile%.txt_1
    end
    &sv tilefull [read %fileunit% readstat]
&end

/* closes file
&if [close %fileunit%] ne 0 &then
    &return Cannot close file

/* deletes file
&sv delstat [delete last.list -file]
&return

```

APPENDIX E: Virtual Deforestation and Aggregation AMLs

These are the most time consuming parts of the process. VDF, DESPIKE3 and ELAPSEDTIME are all used by running VDF. These are the deforestation processes. There are four variable of interest that can be set in the VDF script. These are DZ1, DZ2, POINTSLOST TOLERANCE, AND MAXITERATIONS which are all set in VDF. DZ1 is the distance that a point can be above the local average tin height before it is removed. DZ2 is the maximum distance that a point can be below the average tin height before it is removed, POINTSLOSTTOLERANCE is the percent of points that can be lost in a round before the program jumps out, used to control small data sets. The MAXITERATIONS is the number of rounds that can be run before the program will automatically jump out even if the POINTSLOSTTOLERANCE is not reached. In the DESPIKE3 script there is one setting to note and that is the 2' grid sizes in the tinlattice command.

The TIN2DEM.AML is the final script in the set. It was created to convert all of the deforested tins in the ground folder into one single tin which in turn is used to create a grid and DEM file. The output files are called MASTER_TIN, MASTER_GRD, and MASTER_DEM. No arguments are needed.

VDF.aml

```

/* AML vdf
/* 27 July 2000
/* Ralph Haugerud, USGS - Seattle

/* TILENAME is input TIN, presumably the last-return surface

/* Following segment was added by Matthew Walsh
/* 13 March 2003
/* Allowed the automation of many tin files in one directory

&sv file [filelist * tin.list -tin]

/* opens file list
&sv fileunit [open tin.list openstat -read]
&if %openstat% ne 0 &then

```

```

&return Problem opening file

/* reads file list
&sv tilename [read %fileunit% readstat]
&if %readstat% ne 0 &then
  &return Problem reading file

&do &while %readstat% = 0

/*
/* This is the beginning of the original file from Ralph for processing
/*
&if [null %tilename%] &then &do
  &type USAGE &r vdf <tin>, where <tin> is a last-return
  &type surface of LIDAR data
  &type
  &return
&end

/* units are feet, for data with circa 3' point spacing,
/* and with 2-ft cells in despiking3.aml
&s ProximityTolerance = 0.01
&s dz1 = 0.7
&s dZ2 = 100
&s PointsLostTolerance = 0.1
&s MaxIterations = 16

&s NIterations = 1
&s starttime = [date -time]

/* first pass, no pit elimination
&s teststring = spot < MeanZ + %dZ1%
&s itstart = [date -time]
&r despiking3 %tilename% 0.01 [quote %teststring%]
&r elapsedtime %itstart%
&type Tile %tilename%, time for iteration %NIterations% = %.ElapsedTime%

/* Now, eliminating pits, iterate to quasi-convergence
&s teststring = spot < meanZ + %dZ1% AND spot > meanZ - %dZ2%
&do &until %.PercentLost% < %PointsLostTolerance% OR %NIterations% =
%MaxIterations%
  &s NIterations = [calc %NIterations% + 1]
  &s itstart = [date -time]
  &type Starting iteration %NIterations% at %itstart%
  &r despiking3 %tilename% %ProximityTolerance% [quote %teststring%]

```

```

&r elapsedtime %itstart%
&type Tile %tilename%, time for iteration %NIterations% = %.ElapsedTime%
&end

```

```

&r elapsedtime %starttime%
&type
&type Tile %tilename%
&type Iterations = %NIterations%
&Type Total time = %.ElapsedTime%
/*
/* This is the end of the original file from Ralph for processing
/*
&sv tilename [read %fileunit% readstat]
&end

/* closes file
&if [close %fileunit%] ne 0 &then
&return Cannot close file

/* deletes file
&sv delstat [delete tin.list -file]

&return

```

DESPIKE3.aml

```

/* DESPIKE3.AML
/* AML to take spikes off of a TIN
/* version of 27 July 2000
/* Ralph Haugerud, USGS - Seattle
/* rhaugerud@usgs.gov
/*
&args intin ProximityTolerance teststring

&if [null %intin%] &then &do
&type 'USAGE &R despiked3 <intin> {ProximityTolerance} {Teststring}
&type default ProximityTolerance is 0.01
&type default Teststring is 'spot <= meanZ + 0.7'
&type
&return
&end

/*** set defaults
&if [null %ProximityTolerance%] &then &s ProximityTolerance = 0.01
&if [null %teststring%] &then &s teststring = 'spot <= meanZ + 0.7'

```

```

&s teststring = [translate %teststring%]
&s teststring = [unquote %teststring%]

/**** Turn TIN into point cover
&if [exist xxpoint -cover] &then kill xxpoint all
tinarc %intin% xxpoint POINT

/**** get "mean" value of Z around each point
&if [exist xxmean -grid] &then kill xxmean all
tinlattice %intin% xxmean
[unquote "]
[unquote "]
[unquote "]
2
&if [exist xxmeang -grid] &then kill xxmeang all
grid
xxmeang = focalmean(xxmean)
quit
kill xxmean all
latticespot xxmeang xxpoint meanZ
kill xxmeang all

/**** have had some problems with CREATETIN crashing when it is
/**** presented with NODATA values in meanZ, so get rid of them
&if [exist xxpoint2 -cover] &then kill xxpoint2 all
rename xxpoint xxpoint2
ap
reselect xxpoint2 point meanZ <> -9999
&s trash = [delete x.sl -file]
&s trash = [delete x.sx -file]
writesel x.sl
quit
reselect xxpoint2 xxpoint point x.sl
kill xxpoint2 all

/* create new tin from original points, filtered
/* by criteria in test string
&if [exist %intin% -tin] &then kill %intin% all
createtin %intin% 0 %ProximityTolerance%
cover xxpoint POINT spot 1 # %teststring%
END

&describe xxpoint
&s np1 = %DSC$POINTS%

```

```

&describe %intin%
&s np2 = %TIN$NNODES%

&s .PercentLost = [calc ( %np1% - %np2% ) * 100 / %np1% ]
&type
&type Percent points removed this iteration = %.PercentLost%
&type

kill xxpoint all

&return

```

ELAPSEDTIME.aml

```

/* AML elapsedtime
/* note that starttime = [date -TIME]
&arg starttime

&s endtime = [date -time]
&s sthr = [before %starttime% .]
&s stmin = [after %starttime% .]
&s stsec = [after %stmin% .]
&s stmin = [before %stmin% .]
&s endhr = [before %endtime% .]
&s endmin = [after %endtime% .]
&s endsec = [after %endmin% .]
&s endmin = [before %endmin% .]
&s endtime = [calc %endhr% + ( %endmin% / 60 ) + ( %endsec% / 3600 ) ]
&s starttime = [calc %sthr% + ( %stmin% / 60 ) + ( %stsec% / 3600 ) ]
&s eltime = [calc %endtime% - %starttime% ]
&if %eltime% < 0 &then &s eltime = [calc %eltime% + 24 ]
&s elhr = [trunc %eltime%]
&if %elhr% < 10 &then &s elhr = 0%elhr%
&s eltime = [calc ( %eltime% - %elhr% ) * 60 ]
&s elmin = [trunc %eltime%]
&if %elmin% < 10 &then &s elmin = 0%elmin%
&s elsec = [round [calc ( %eltime% - %elmin% ) * 60 ] ]
&if %elsec% < 10 &then &s elsec = 0%elsec%
&s .ElapsedTime = %elhr%:%elmin%:%elsec%
&return

```

APPENDIX F: Programs Used For Flow Direction

SAGA – <http://geosun1.uni-geog.gwdg.de/saga/>

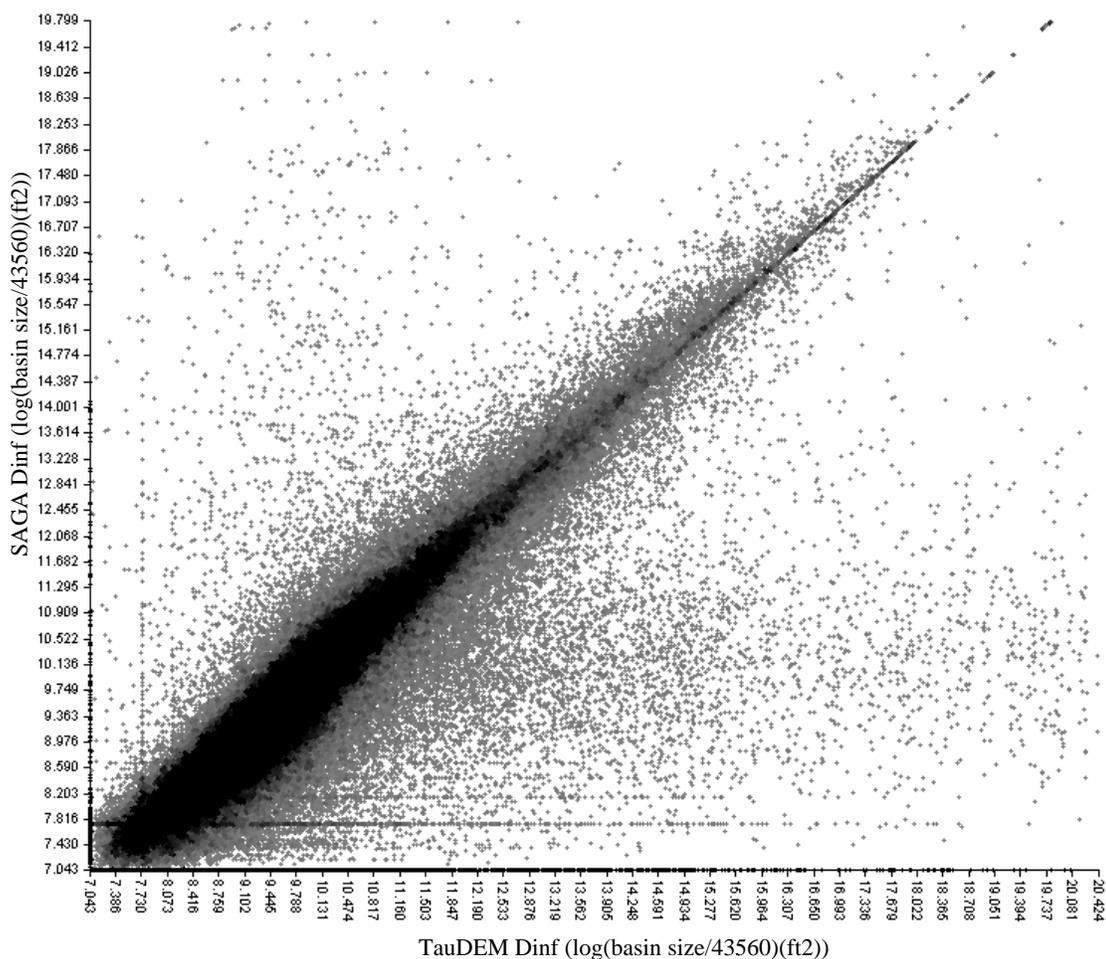
ArcGIS – <http://www.esri.com/>

TauDEM – <http://hydrology.neng.usu.edu/taudem/>

TAS – <http://publish.uwo.ca/~icreed/tas.htm>

SAGA performed the D8, Dinf, MFD, and DEMON flow direction algorithms that were tested of this study. The main reasoning in using SAGA was that most programs could not handle a large LiDAR dataset. TAS, TauDEM and at times ArcGIS would fail when attempting to use a 2-m LiDAR DEM. SAGA used a memory cache to process data. In doing this, process time increased significantly in calculating DEMON, Dinf, and MFD flow algorithms. Demon at 2-m grid size took at around a day to compute while MFD and Dinf was half that. D8 only took a few hours. Using 6-m grid cell size, process time for all flow direction algorithms decreased to a few hours.

SAGA was verified using ArcGIS for the D8, TauDEM for Dinf and TAS for MFD on a 10-m DEM. DEMON was verified by Dr. Steven Burges. Regression analysis was performed on the flow accumulation outputs from the 10-m DEM by the flow direction types. The only differences were in how the algorithms handled flat terrain. Appendix Figure F1 and Appendix Table F1 is an example of the regression analysis that was done to verify SAGA performed the flow direct algorithms accurately.



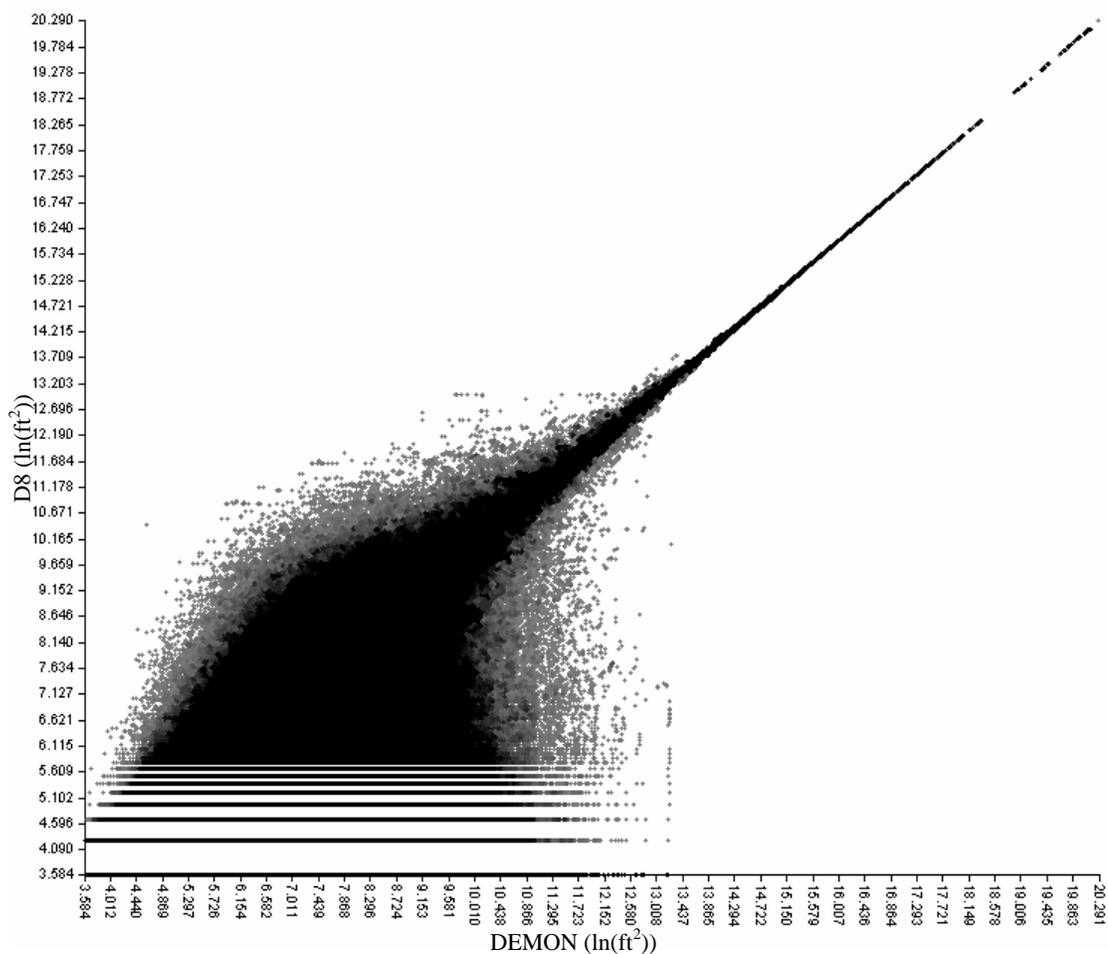
Appendix Figure F1. Correlation between Dinf algorithms by TauDEM and SAGA

Appendix Table F1. Regression outputs from Appendix Figure E1

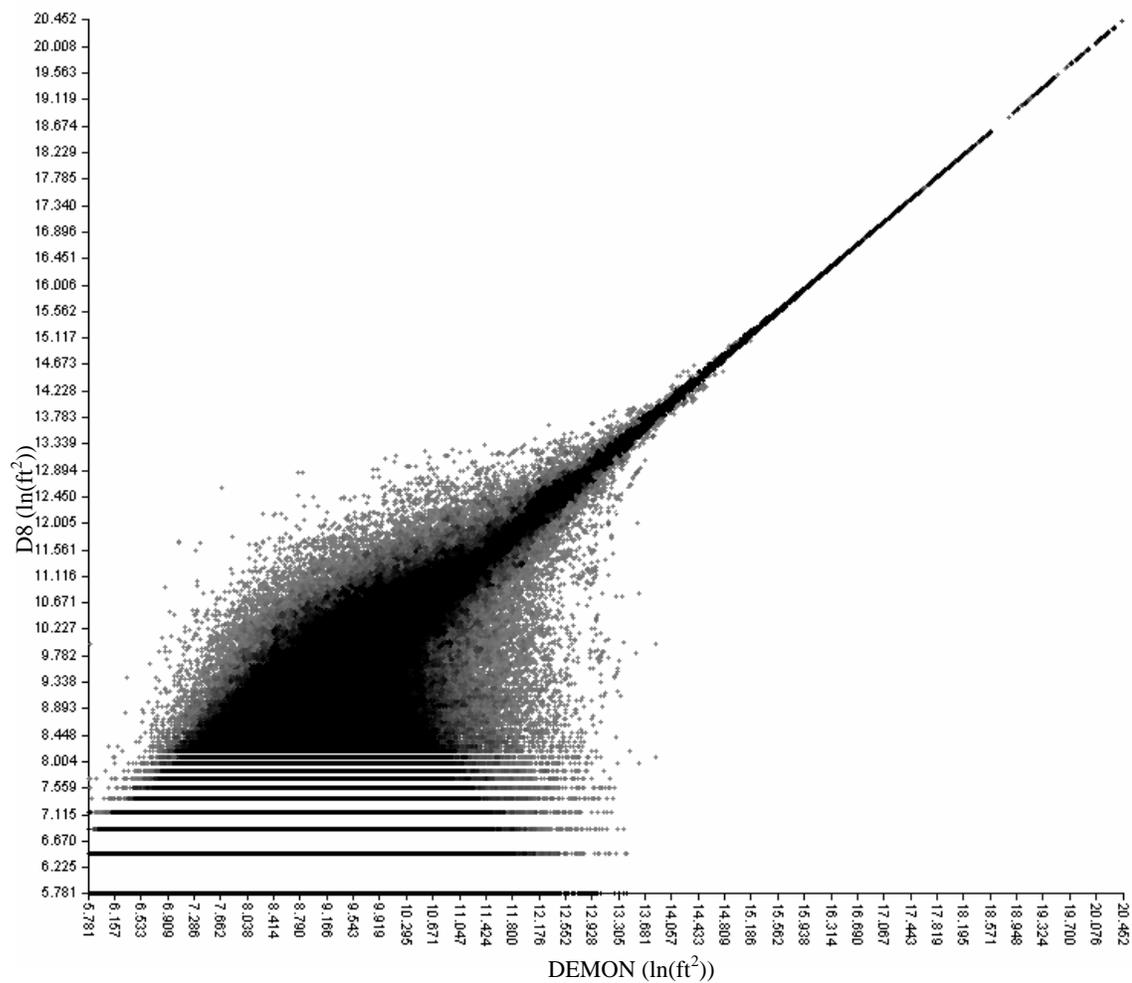
<u>Regression Analysis</u>		
N = 1966041	Linear Regression:	$Y = 0.861927 * X + 1.228089$
SAGA dinf:		($r=0.8654, r^2=0.7489$)
Min. = 7.043411	Max. = 20.423664	
Arithmetic Mean = 9.711222		
Variance = 1.725083		
Standard Deviation = 1.313424		
TauDEM dinf:		
Min. = 7.043411	Max. = 19.798889	
Arithmetic Mean = 9.584567		
Variance = 1.711332		
Standard Deviation = 1.308179		

APPENDIX G: Regression Correlation of Flow Direction Algorithms with Increased Resolution

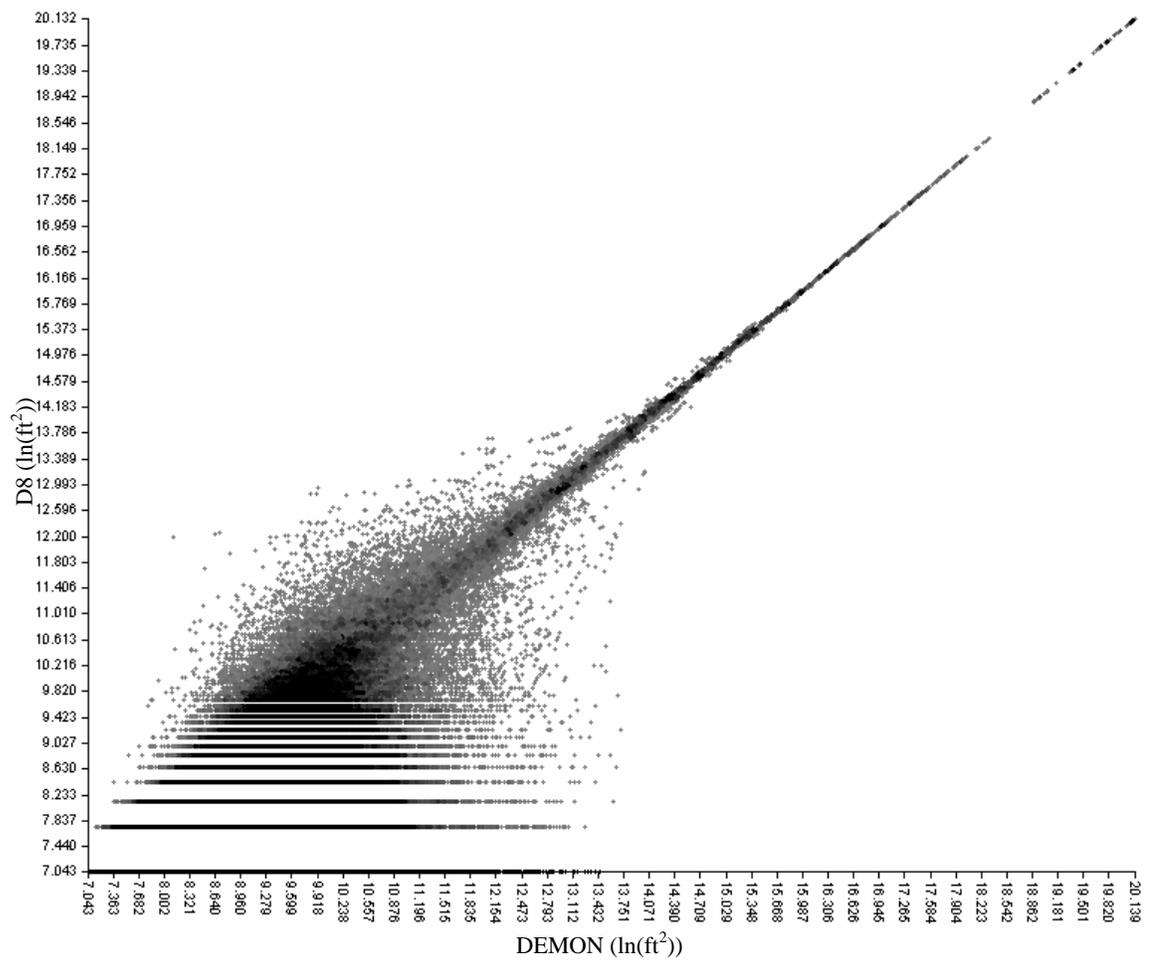
D8 vs. DEMON



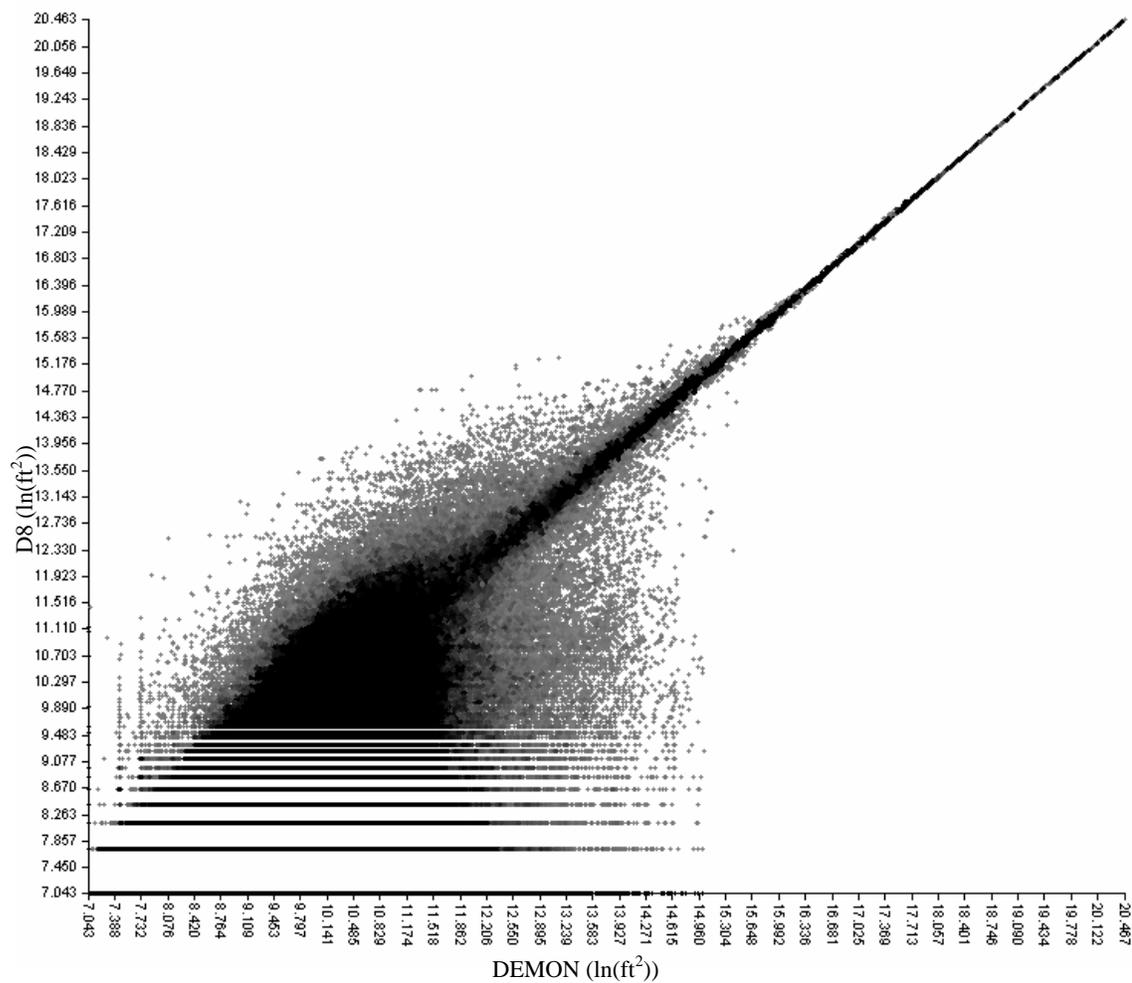
Appendix Figure G1. Plot of entire catchment area for a 2-m LiDAR DEM at Tahoma State Forest (Natural Log Values). The correlation between D8 and DEMON flow direction models indicates that D8 is a sufficient model for determining stream networks.



Appendix Figure G2. 6-m LiDAR Plot of catchment area comparing Demon and D8

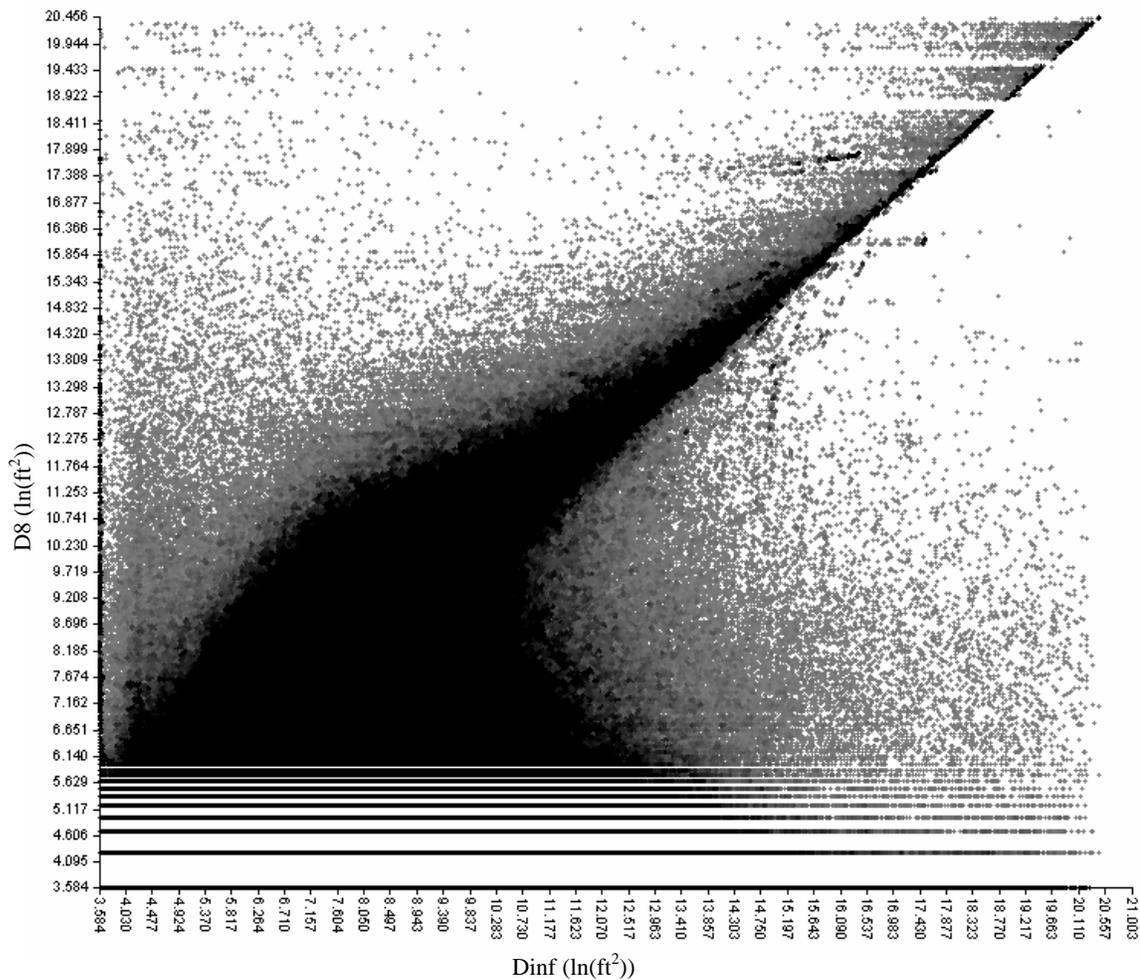


Appendix Figure G3. 10-m LiDAR Plot of catchment area comparing Demon and D8

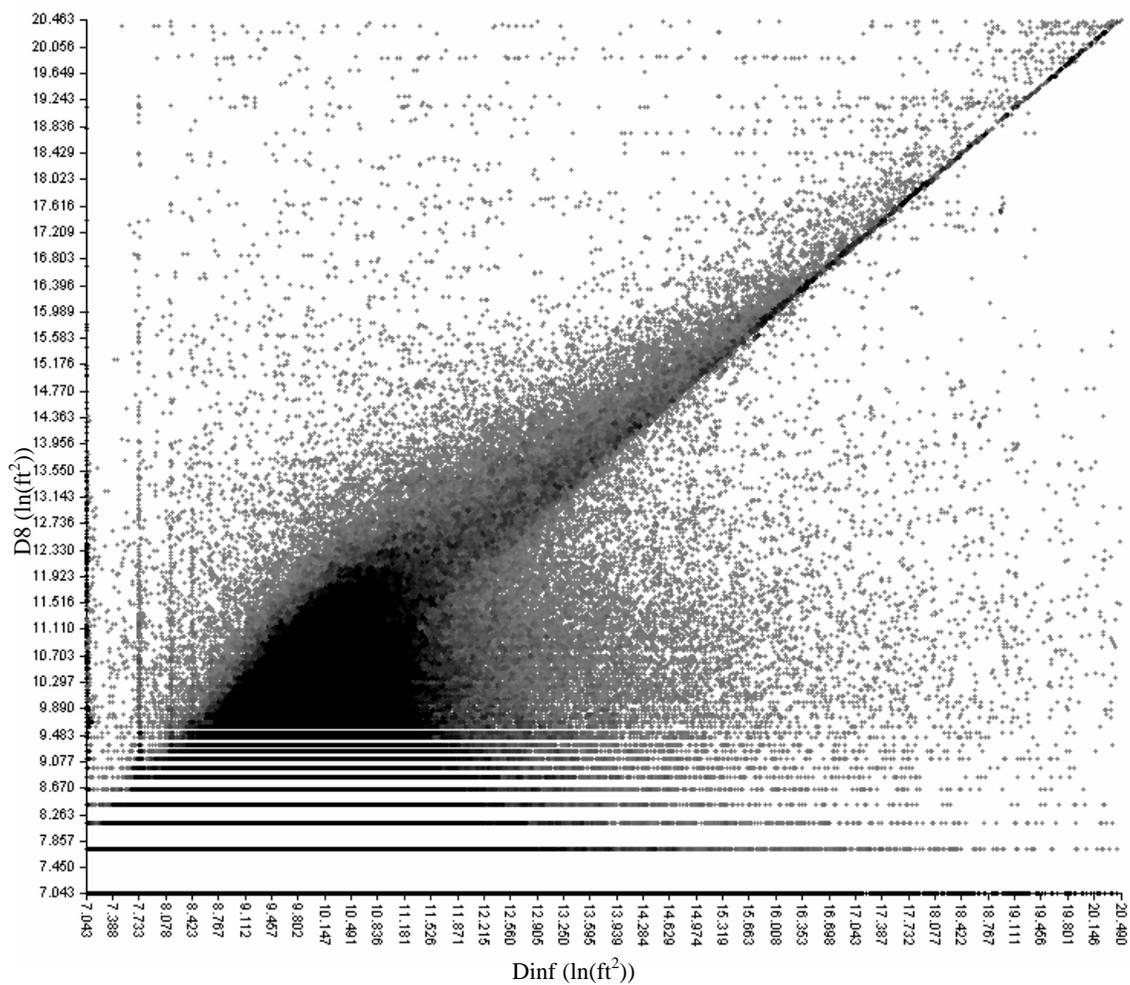


Appendix Figure G4. 10-m USGS plot of catchment area comparing Demon and D8

D8 vs. Dinf

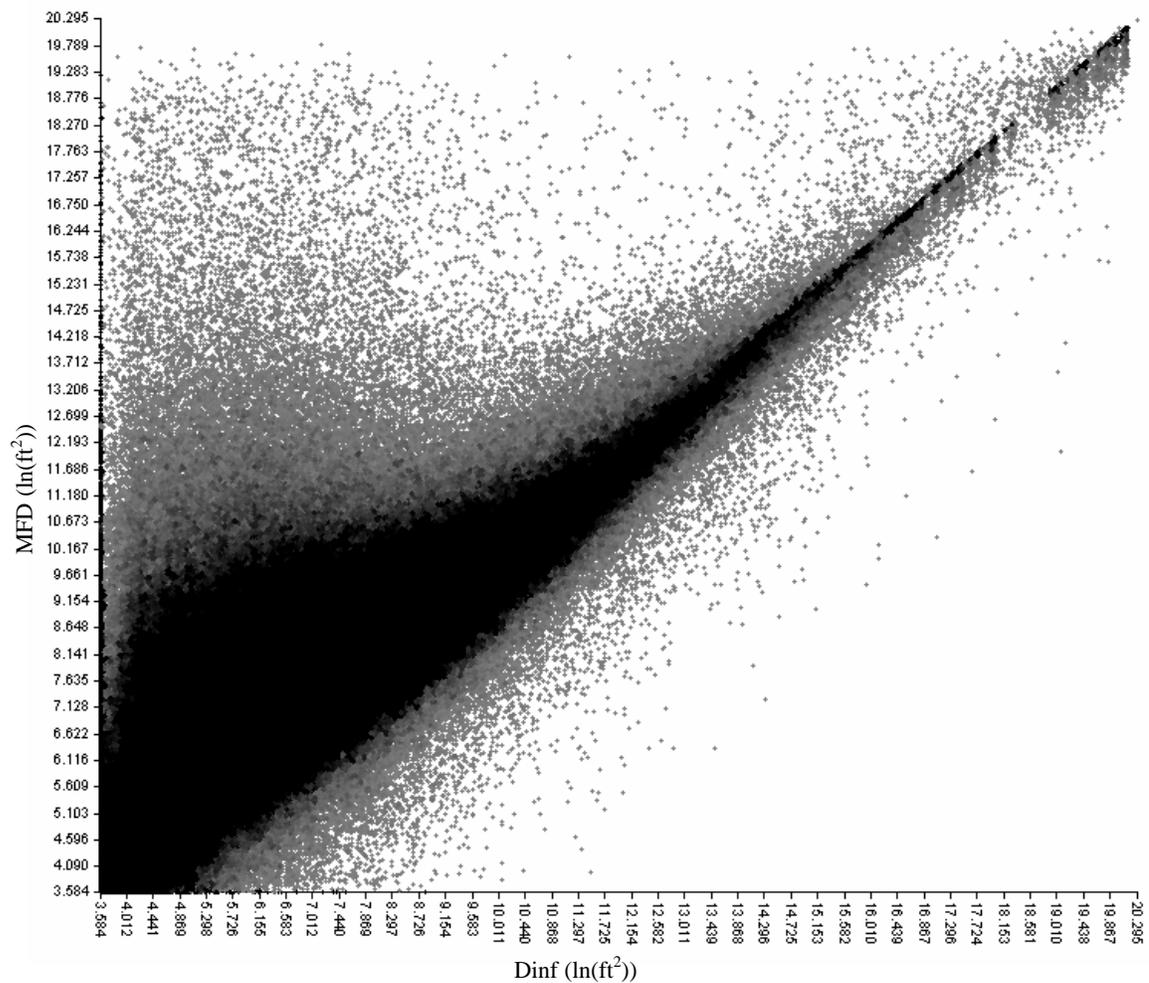


Appendix Figure G5. 2-m LiDAR Plot of catchment area comparing Dinf and D8

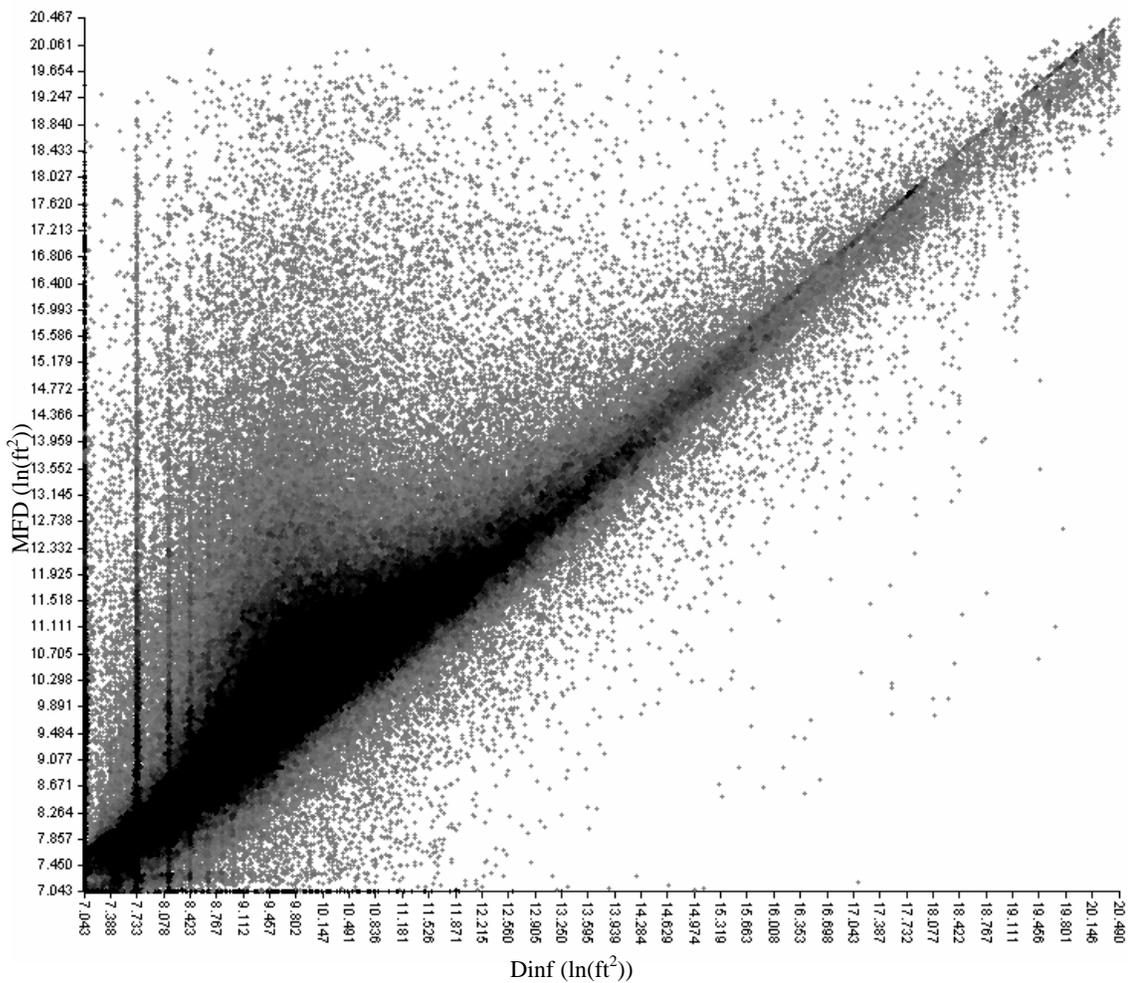


Appendix Figure G6. 10-m USGS Plot of catchment area comparing Dinf and D8

MFD vs. Dinf



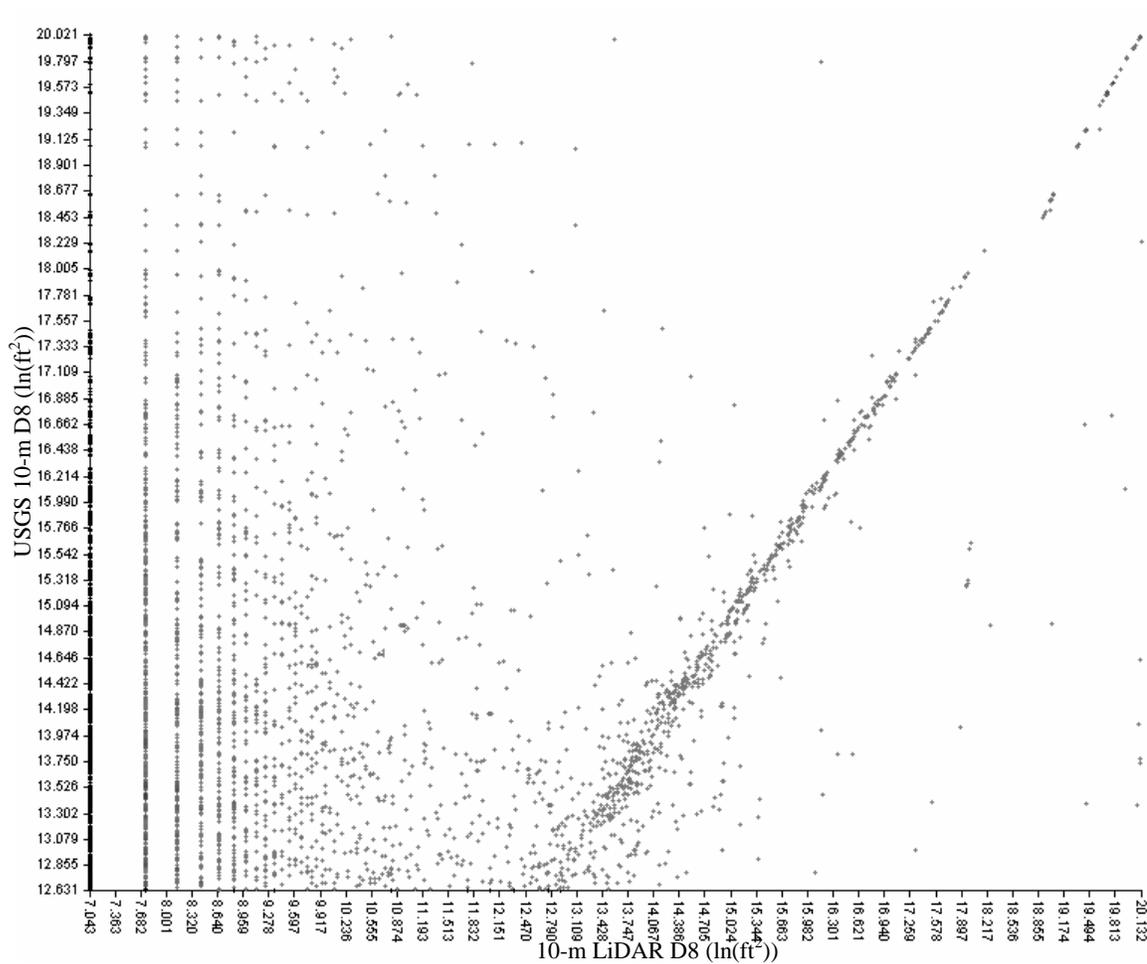
Appendix Figure G7. 6-m LiDAR Plot of catchment area comparing Dinf and MFD



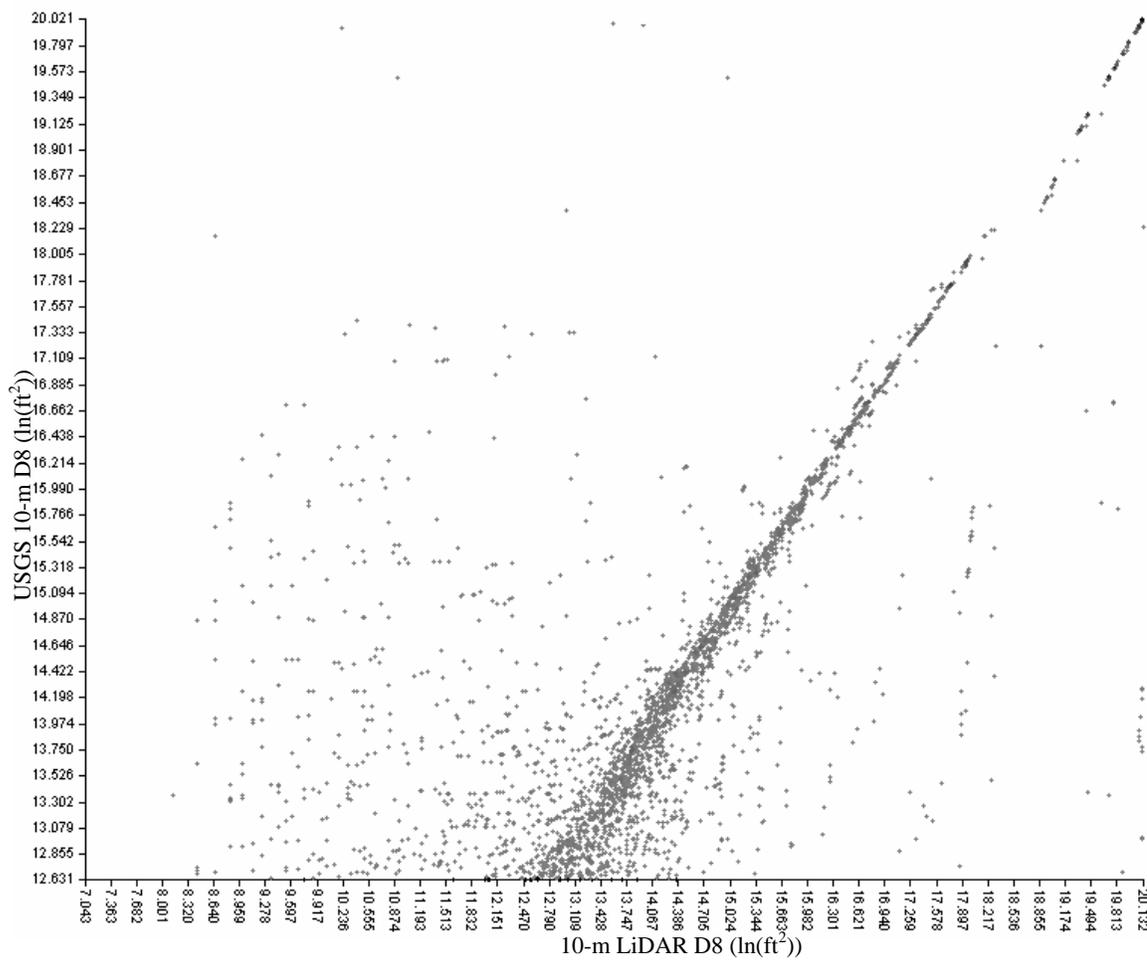
Appendix Figure G8. 10-m USGS Plot of catchment area comparing Dinf and MFD

APPENDIX H: D8 Comparison at 10-m Resolution of USGS and LiDAR Generated

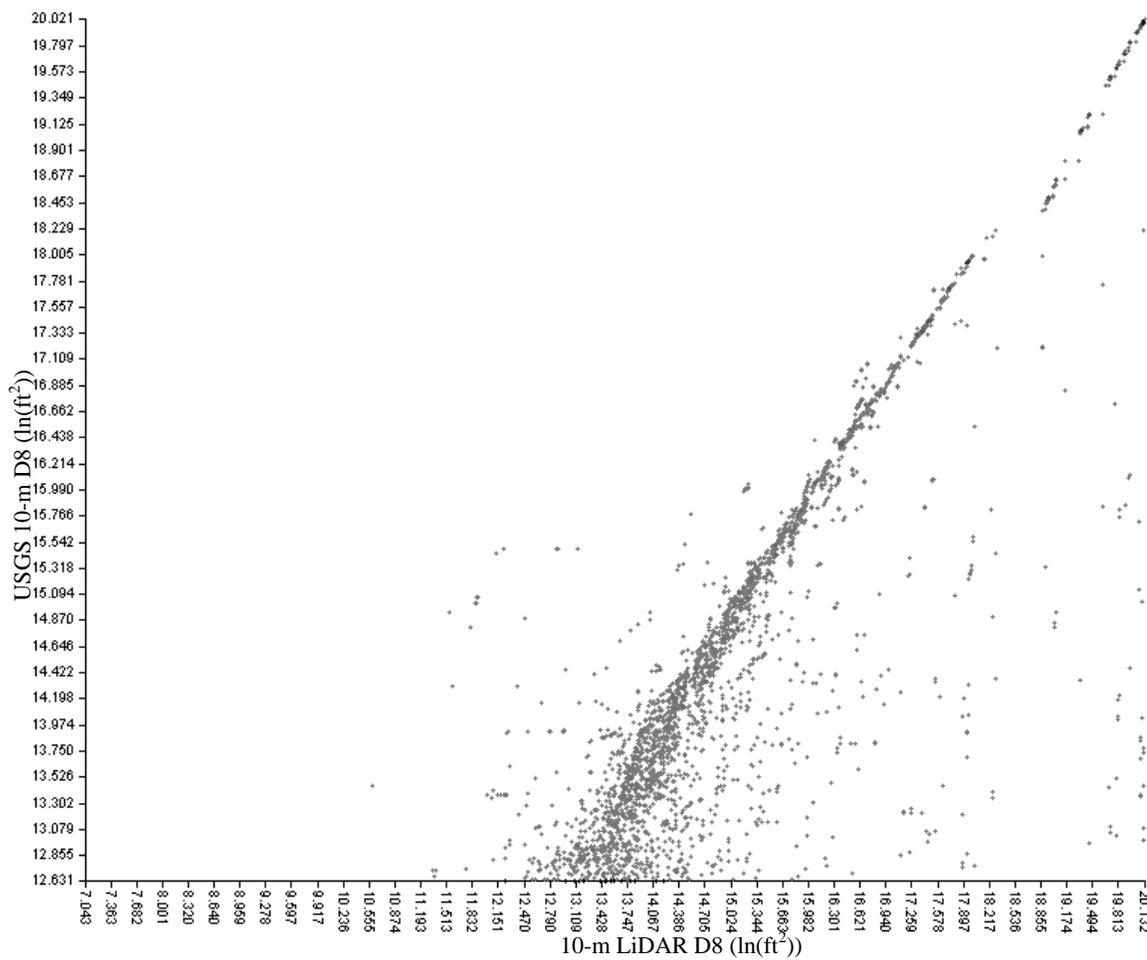
Data



Appendix Figure H1. The D8 flow algorithm applied to the USGS and LiDAR Generated 10-m DEM. Illustrates that stream channels with a catchment size of about 12-ha and greater somewhat converge between the two DEM's with regards to D8. When catchments are less than 12-ha, differences in stream channel location are shown.



Appendix Figure H2. USGS and LiDAR Generated 10-m, D8 Comparison. Streams identified as identical if 30-m apart.



Appendix Figure H3. USGS and LiDAR Generated 10-m, D8 Comparison. Streams identified as identical if 90-ft apart.

APPENDIX I: Statistics For Binary Linear Regression

The statistics were broken into steps to determine which variables would best work for an equation. All statistics were done in SPSS. Below are the significant steps taken and the variables in each step. Step 3 is the most significant.

Step 1: Log10(Basize)

Step 2: Log10(Basize), Precipitation

Step 3: Log10(Basize), Precipitation, Slope

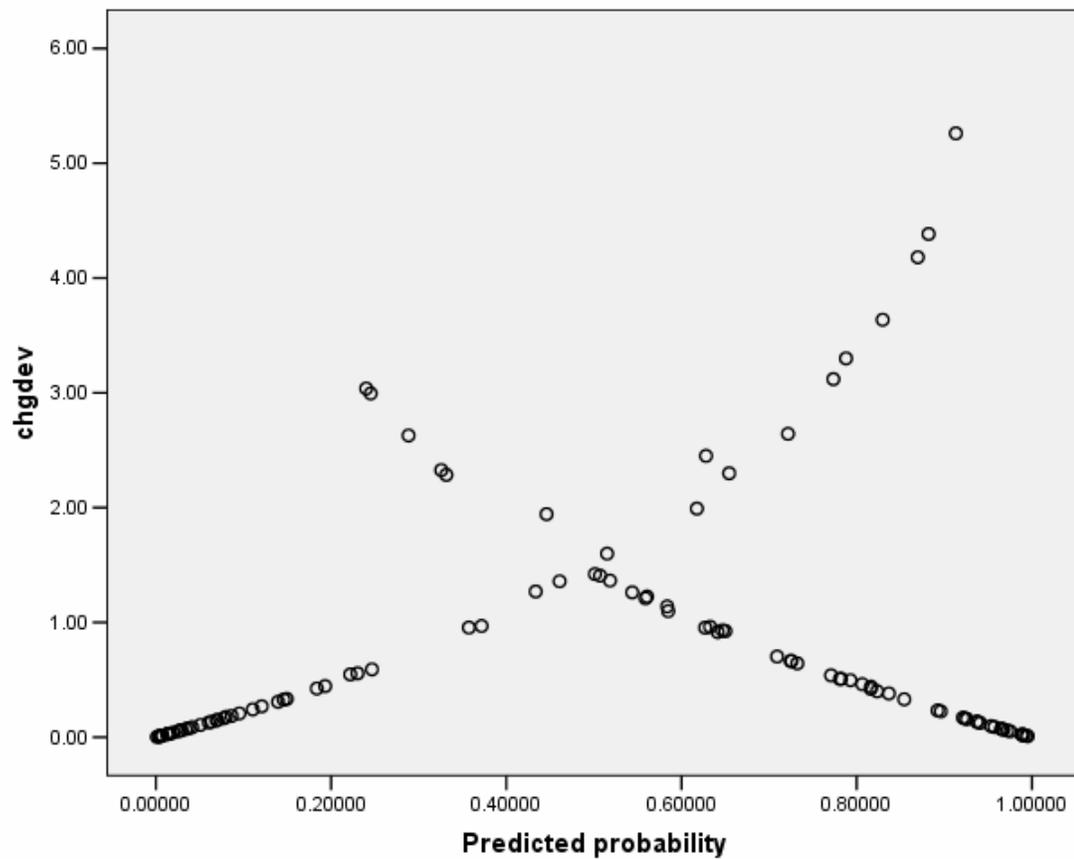
Elevation, downstream gradient and site class are not part of the steps due to it being insignificant for the model.

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	8.864	7	0.263
2	9.900	8	0.272
3	10.262	8	0.247

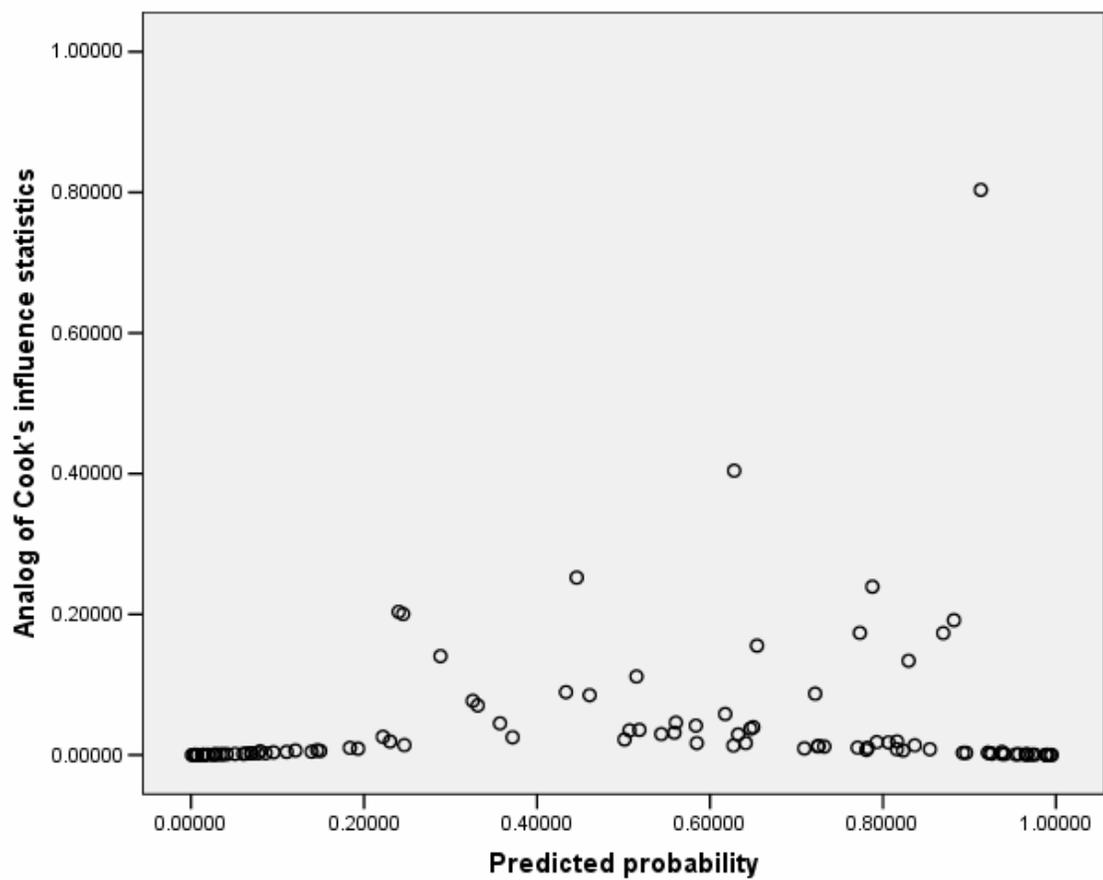
Appendix Table II. The Hosmer-Lemeshow statistic. Indicates a poor fit if the significance value is less than 0.05.

Contingency Table for Hosmer and Lemeshow Test						
		Head = .00		Head = 1.00		Total
		Observed	Expected	Observed	Expected	
Step 1	1	14	13.289	0	0.711	14
	2	10	9.966	1	1.034	11
	3	8	9.082	3	1.918	11
	4	8	7.142	3	3.858	11
	5	3	5.195	8	5.805	11
	6	4	3.662	7	7.338	11
	7	2	2.283	9	8.717	11
	8	4	1.572	7	9.428	11
	9	0	0.810	15	14.190	15
Step 2	1	11	10.632	0	0.368	11
	2	11	10.225	0	0.775	11
	3	10	9.602	1	1.398	11
	4	6	7.908	5	3.092	11
	5	6	5.435	5	5.565	11
	6	2	4.162	9	6.838	11
	7	2	2.679	9	8.321	11
	8	4	1.567	7	9.433	11
	9	1	0.644	10	10.356	11
	10	0	0.147	7	6.853	7
Step 3	1	11	10.903	0	0.097	11
	2	11	10.543	0	0.457	11
	3	11	9.741	0	1.259	11
	4	6	7.948	5	3.052	11
	5	4	5.371	7	5.629	11
	6	4	3.887	7	7.113	11
	7	2	2.508	9	8.492	11
	8	4	1.466	7	9.534	11
	9	0	0.554	11	10.446	11
	10	0	0.079	7	6.921	7

Appendix Table I2. This statistic is the most reliable test of model fit for SPSS binary logistic regression, because it aggregates the observations into groups of cases.



Appendix Figure I1. Deviance plot change helps identify cases that are poorly fit by the model.



Appendix Figure I2. The shape of the Cook's distances plot generally follows that of the previous figure, with some minor exceptions. These exceptions are high-leverage points, and can be influential to the analysis.

Variables not in the Equation					
			Score	df	Sig.
Step 1	Variables	PRECIP	4.452	1	0.035
		Slope	3.242	1	0.072
		Elevation	2.268	1	0.132
		Downgrad	0.611	1	0.434
		Site Class	0.152	1	0.696
	Overall Statistics		11.696	5	0.039
Step 2	Variables	Slope	6.284	1	0.012
		Elevation	0.450	1	0.502
		Downgrad	2.153	1	0.142
		Site Class	0.148	1	0.700
	Overall Statistics		7.889	4	0.096
Step 3	Variables	Elevation	0.644	1	0.422
		Downgrad	1.126	1	0.289
		Site Class	0.121	1	0.728
	Overall Statistics		1.658	3	0.646

Appendix Table I3. Forward stepwise methods variables left from steps

Model if Term Removed					
Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1	Log10(BASIZE)	-73.474	55.107	1	0.000
Step 2	Log10(BASIZE)	-73.455	59.991	1	0.000
	PRECIP	-45.920	4.921	1	0.027
Step 3	Log10(BASIZE)	-73.422	66.715	1	0.000
	PRECIP	-44.246	8.363	1	0.004
	Slope	-43.460	6.789	1	0.009

Appendix Table I4. The variables chosen by the forward stepwise method all having significant changes in -2 log-likelihood.

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	91.840	0.405	0.541
2	86.919	0.432	0.577
3	80.130	0.468	0.623

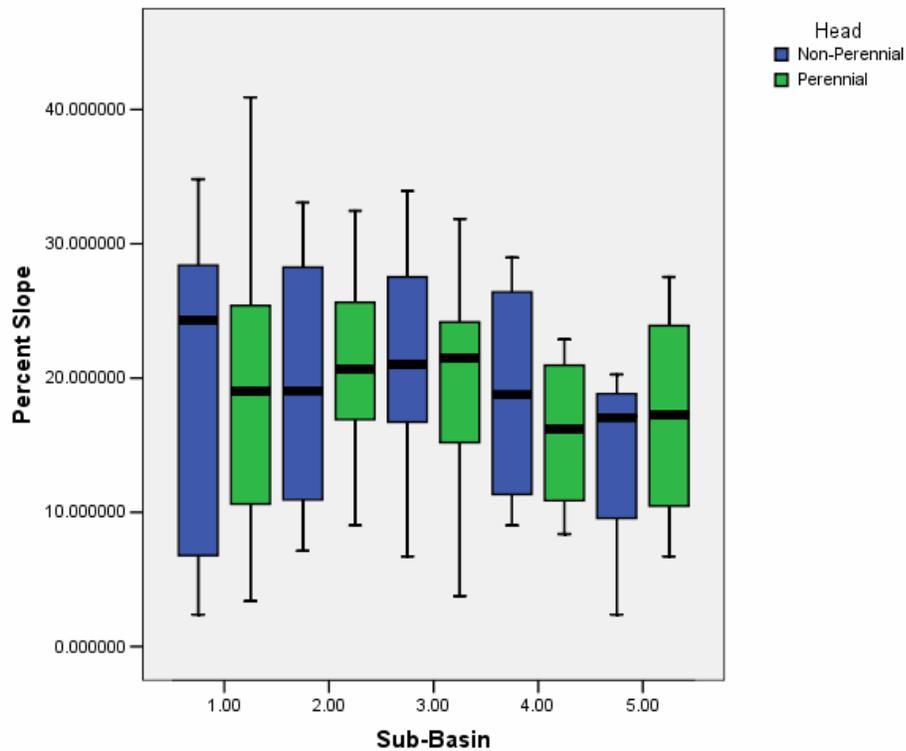
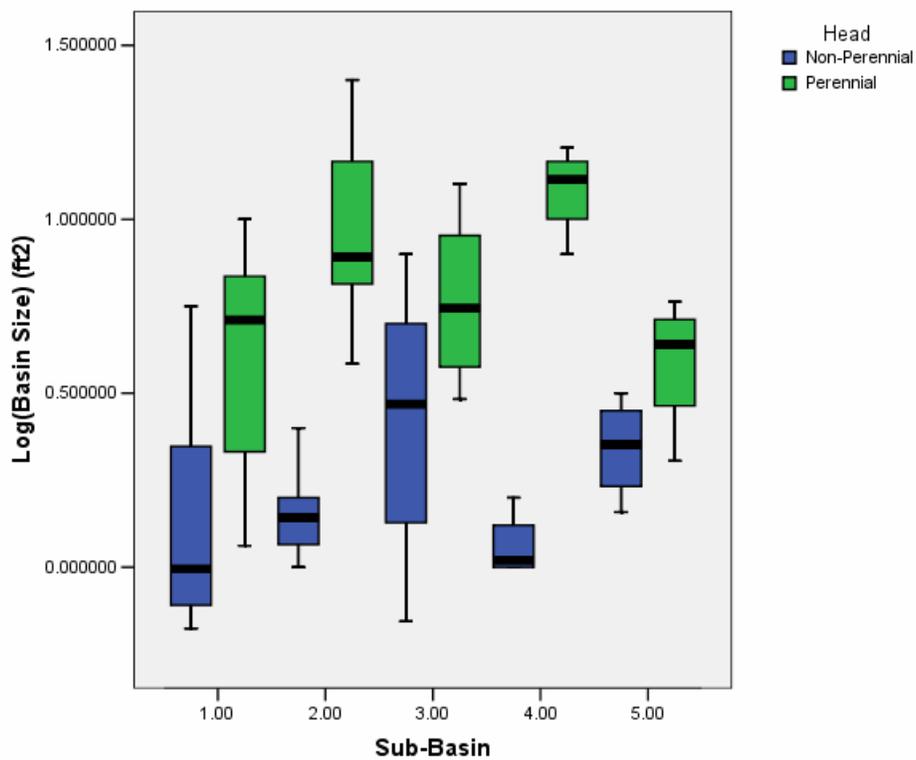
Appendix Table I5. The pseudo r-squared statistics.

Classification Table(a)						
Observed			Predicted			
			Head		Percentage Correct	
			0	1		
Step 1	Head	0	41	12	77.358	
		1	8	45	84.906	
Overall Percentage			81.132			
Step 2	Head	0	40	13	75.472	
		1	9	44	83.019	
Overall Percentage			79.245			
Step 3	Head	0	41	12	77.358	
		1	6	47	88.679	
Overall Percentage			83.019			

Appendix Table I6. The classification table indicating the practical results of using the logistic regression model.

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
Step 1(a)	Log10(BASIZE)	5.115	0.935	29.948	1	0.000	166.464	26.653	1039.659
	Constant	-2.711	0.570	22.617	1	0.000	0.066		
Step 2(b)	Log10(BASIZE)	5.753	1.064	29.241	1	0.000	315.028	39.157	2534.451
	PRECIP	0.336	0.165	4.157	1	0.041	1.400	1.013	1.934
	Constant	-30.754	13.849	4.931	1	0.026	0.000		
Step 3(c)	Log10(BASIZE)	7.235	1.425	25.766	1	0.000	1386.737	84.879	22656.314
	PRECIP	0.477	0.184	6.697	1	0.010	1.612	1.123	2.313
	SLOPE	0.096	0.040	5.786	1	0.016	1.101	1.018	1.191
	Constant	-45.172	16.050	7.921	1	0.005	0.000		

Appendix Table I7. The parameter estimates table summarizing the effect of each predictor.



Appendix Figure I3. Boxplots comparing the distribution of % slope and basin size values for PIP's.