

**CE 394K.3 GIS in Water Resources**

# **GIS feature extraction tools in diverse landscapes**

---

**Final Project**

**Anna G. Kladzyk  
M.S. Candidate, Expected 2015  
Department of Environmental and Water Resources Engineering  
12/6/2013  
University of Texas at Austin**

## 1. Introduction

By observing landscapes over time and space, signatures of the hydrologic connectivity emerge. For water resources management, observations of drainage network structure and changes therein, provide critical insights into river basin morphology. Parameters that define drainage basins include terrain slope and curvature, channel density and character, as well as geologic and vegetative features. The identification of land surface features illuminates patterns signifying advancing mass and energy across the terrain. Improved understanding of how these patterns manifest as drainage basins is a key aspect to advancing the design and implementation of infrastructure, restoring and protecting vulnerable ecosystems, and preparing for and mitigating extreme climate events.

The advent of high-resolution imagery data, obtained from advanced scanning technologies, engendered new capabilities with which to observe topographic features at a landform process-scale [2]. Signatures of hydrologic and geomorphic patterns can be extracted from Light Detection and Ranging (LIDAR) imagery data using computational algorithms. Examples of extracted features using current technologies include: channel networks, terrain slope and curvature, landslide scars, and engineered features such as roads and ditches. However, limitations in feature extraction exist, resulting from the LIDAR data itself and the analysis processes. For example, current extraction methods are less robust in low-relief and vegetative terrains, in part due to necessary filtering methods which are used to eliminate vegetation or “noise” in order to obtain bare earth terrain. These filtering operations may inadvertently eliminate or distort important landscape features. Feature extraction tools also perform differently when analyzing basins of low-relief terrain versus those in mountainous regions with diverse elevation gradients. Analysis of high-resolution imagery data provides terrain data at a resolution appropriate for interpreting information about surface processes and hydrologic connectivity, but evaluation of extraction tool performance is necessary to develop more robust analysis frameworks for diverse conditions and regions.

In the analysis of watershed landscapes, important parameters describing the characteristics of the basin of interest can be derived using algorithms to perform feature extraction on high resolution imagery data. Three of these parameters are: (1) slope of the terrain, (2) curvature of the terrain, and (3) accumulation area, or the area from which water flows into a single depression. Furthermore, the graph of the accumulation area (A) versus slope (S) is used to estimate a fourth critical characteristic of the basin; the initiation area or channel head. This parameter corresponds to the inflection point on the A v. S graph. Once the channel head is identified, it is possible to generate the channel network for the basin. Different feature extraction frameworks generate these parameters with varying methods. The comparison of two calculation methods in two physically diverse basins is the subject of this study.

The study focuses on two basins; the first being the Eel River Basin in northern California, which is characterized by highly vegetative land and stark elevation variations. The second is the Le Sueur Basin in southern Minnesota, which has a low-relief terrain and features meandering river channels.

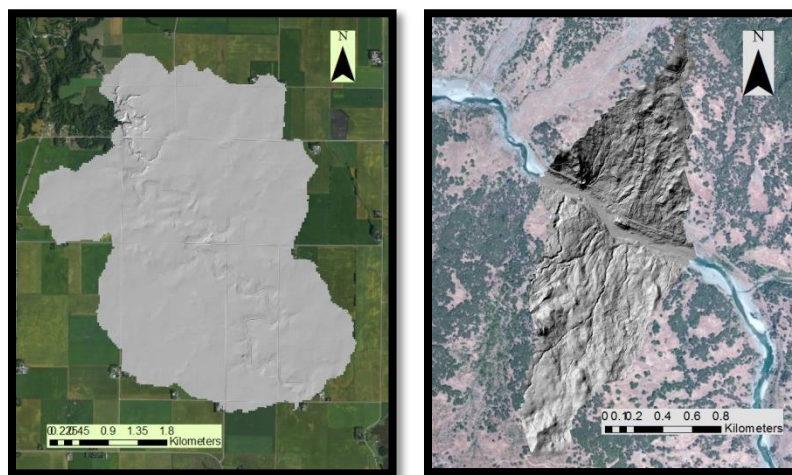


Figure 1. a) Sub-basin in the Le Sueur Basin, MN. b) Sub-basin in the Eel River Basin

Another major difference between the two basins is the Le Seuer has a significant amount of engineered features, including roads and ditches, whereas the landscape of the Eel is predominantly uninfluenced by anthropogenic features. The methods employed to extract features from these two basins include those within the Hydrology tools in the ArcMap program, as well as GeoNet, a MatLab-based computational tool. Using these tools, the following feature information for the sub-basin including: accumulation area, curvatures, and slopes. Statistical and graphical analysis of these basin attributes will illustrate GIS extraction tools' capabilities and limitations in diverse landscapes.

## **2. Objectives**

The first objective of this study is to extract channel network features from each sub-basins using different extraction methods. The differences in the extraction procedures for the frameworks are primarily found in the methods used to filter the digital elevation model; a process required to determine the flow direction across the basin's surface. Otherwise the extraction procedures used by the two different tools are very similar. As stated in the introduction of this paper, the features of interest to be extracted for both basins include:

- i. Slope- the first derivative or gradient of the surface*
- ii. Curvature- the second derivative of the surface (the slope of the slope)*
- iii. Flow accumulation area- the area from which water flows into a single depression*
- iv. Channel network- the drainage network as defined by a single channel head within the basin*

In order to interpret the results of these extraction frameworks, stochastic and graphical methods are be employed to compare the results. To compare the curvature results for the basins, quantile-quantile plots are used to identify the most divergent and convergent features in the basin. Channel heads in the basin can be identified as the inflection point on the slope versus area graph. These plots are developed for both basins and then compared for a discussion on the different frameworks' performance in diverse landscapes.

## **3. Methods**

### **a. Data acquisition**

LIDAR data for the Eel River basin in California was acquired at 1-meter resolution from the Open Topography website (<http://opentopography.org/>). This data was collected in 2006 by Josh Roering of the University of Oregon. LIDAR for the Le Seuer Basin at 3-meter resolution was acquired from a colleague in my research group, Harish Sangireddy, a doctoral candidate at the University of Texas at Austin.

The National Hydrography Dataset (NHD) was acquired from the United States Geological Survey website. This dataset containing shapefiles of hydrologic flow lines and catchments and is used to delineate the sub-basins for each region.

### **b. Processing**

#### **i. ArcMap processing**

The LIDAR data was loaded into ArcMap as geotiff files. The NHD dataset projected into the spatial reference of the lidar data, and then the catchment shapefile was used to extract sub-basins from each dataset. The surface processing tools within the Spatial Analyst tools were then used to begin extracting feature information from the data. First the curvature tool was using to determine the second derivative of the surface. The ArcMap framework performs this by fitting a fourth-order polynomial to the surface.

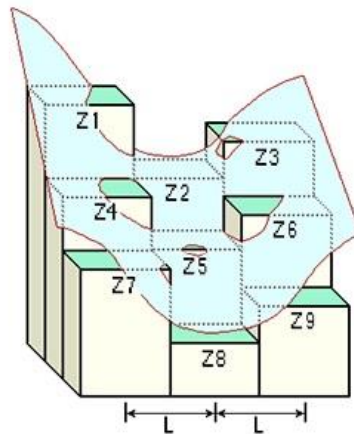


Figure 2. Fitting of a fourth-degree polynomial to the surface.  
Image: ArcGIS Resources; 'How Curvature Works'

The slope of the surface was also found using the slope tool in the Spatial Analyst toolbox. The output of this operation can be either “degree”, which measures the angle of incline of the surface, or “percent rise”, which outputs the incline of the surface in a percent measure. The raster calculator was then used to edit the slope raster so the values would be in fractional percentages. The following simple equation was used for this process, and the output of this process was saved as a new raster file.

$$\text{slope}/100$$

The accumulation area raster was generated through a series of steps in ArcMap. To begin, a buffer of 100 meters was added to the catchment. This is done to prevent any erroneous flow direction values which may occur at the outer edges of the catchment to be processed. After the accumulation areas for the buffered catchment are calculated, this dataset is clipped to the original catchment extent to isolate the data for the basin of interest. The Fill function was used to fill sinks in the raster. Sinks are a common effect of LIDAR data pre-processing used to remove vegetation and obtain the bare earth. These sinks are isolated depressions in which water becomes “trapped”, when performing flow accumulation processing. The Fill function identifies these sinks and fills them to the elevation of the surrounding pixel with the lowest value. Some pits in the data may be natural formations and not the result of pre-processing, therefore this function may remove valuable information about the terrain surface.

The next step is to generate the flow direction rasters for the terrain. The Flow Direction function uses the Fill function output file, or the filled digital elevation model as the input. The outputs of this function are two raster datasets: the flow direction and the percent drop. The percent drop is calculated as the difference in elevation values divided by the length between two pixels from center to center, (ArcGIS Resource Center). The flow direction is determined using the D8 model which assigns one of eight flow directions corresponding to the steepest drop from one cell to the next. The coding for the direction of the flow is shown in the image below.

32	64	128
16		1
8	4	2

Figure 3. D8 flow direction codes.  
Image: ArcGIS Resource Center

Finally the flow accumulation area is calculated using the Flow Accumulation function. This tool uses the flow direction raster as an input and calculates the number of upstream cells that flow into each cell. This flow accumulation raster can then be used to identify the stream pixels in the basin.

**ii. GeoNet processing**

GeoNet performs filtering operations in a different manner than ArcMap. Nonlinear filtering and geodesic minimization principles are used to extract features of the terrain. Nonlinear filtering differs from the pit-filling operations employed in ArcMap in that it minimizes small-scale variability in the LIDAR, while enhancing features of interest such as channel banks [1]. The flow accumulation is calculated using a D8 model, as in ArcMap.

**c. Plots**

**i. Quantile-quantile plots**

Quantile-quantile plots are useful for the identification of channel and ridge features within a basin. The curvature values are ranked and then plotted against the standard normal variate. Curvature values that deviate from the linear relationship on the plot are the most convergent and divergent features. These plots were made in MatLab using exported geotiff files of the curvature rasters.

**ii. Slope versus accumulation area plots**

Plotting the slope versus the accumulation area for each basin is used to identify pixels in the data that correspond to the channel head of the basin. These pixels are identified from the inflection point in the graph. The inflection point in these plots appears as a concave downward inflection. These plots were also made in MatLab using geotiff file versions of the raster files.

**4. Results**

**a. ArcMap processing**

**i. Eel River sub-basin**

The following images depict the raster values of the processing results for the Eel River sub-basin, including: slope, curvature, and accumulation area.

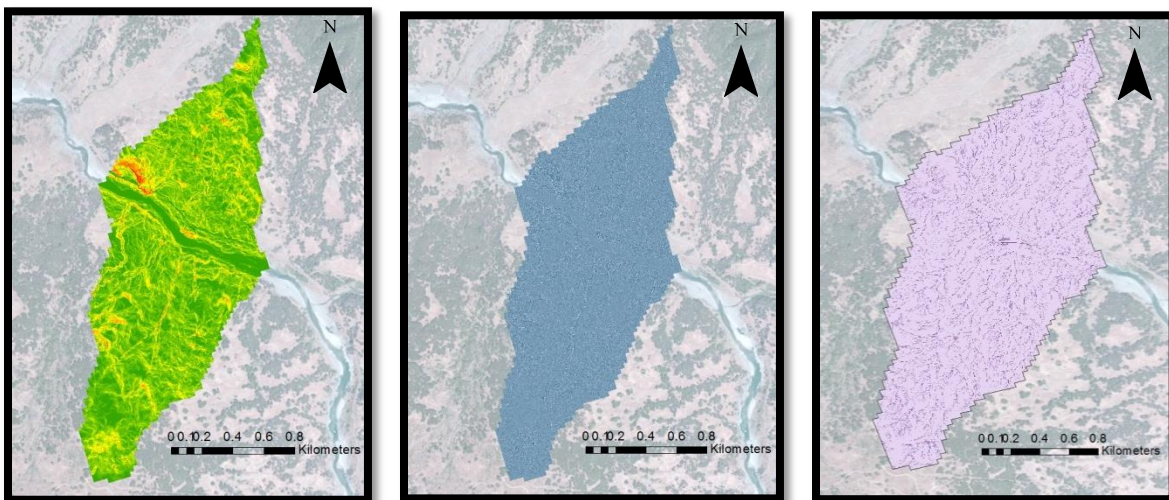


Figure 4. From left to right: slope, curvature, and accumulation area rasters for Eel River sub-basin.

The chart below show the quantile-quantile plot for the curvature values calculated in ArcMap. The deviations from normality indicate the convergent and divergent features. This plot indicates that the likely channelized pixels begin at values of curvature slightly greater than zero.



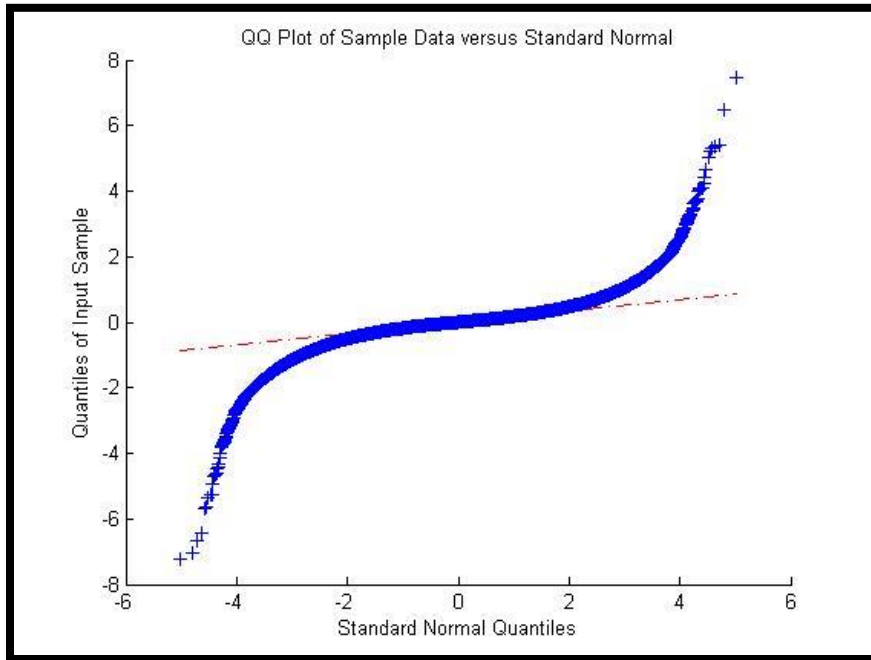


Figure 5. Quantile-quantile plot of curvature values in Eel River sub-basin.

The following plot shows the log-base 10 values of slope versus the log-base 10 values of accumulation area. The inflection point in the chart indicates the pixel values that correspond to the channel head. As shown, very clear inflection is not easily identifiable from this chart, but the value of 3000 sq. meters is used, as there appears to be a slight inflection in that region.

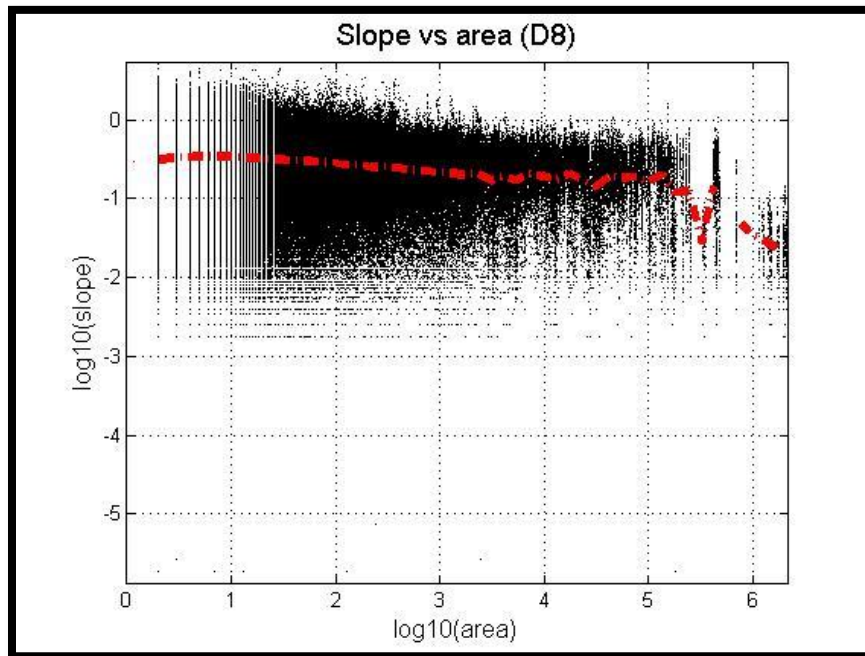


Figure 6. Plot of log-slope versus log-accumulation area. The inflection point is identified as the inflection slightly past  $\log(3)$  on the abscissa axis; corresponding to the accumulation area equal to 3000 sq. meters.

The channel network is extracted using an accumulation area threshold value of 3000 sq. meters. The extracted network is shown over the digital elevation model and the satellite imagery base map.

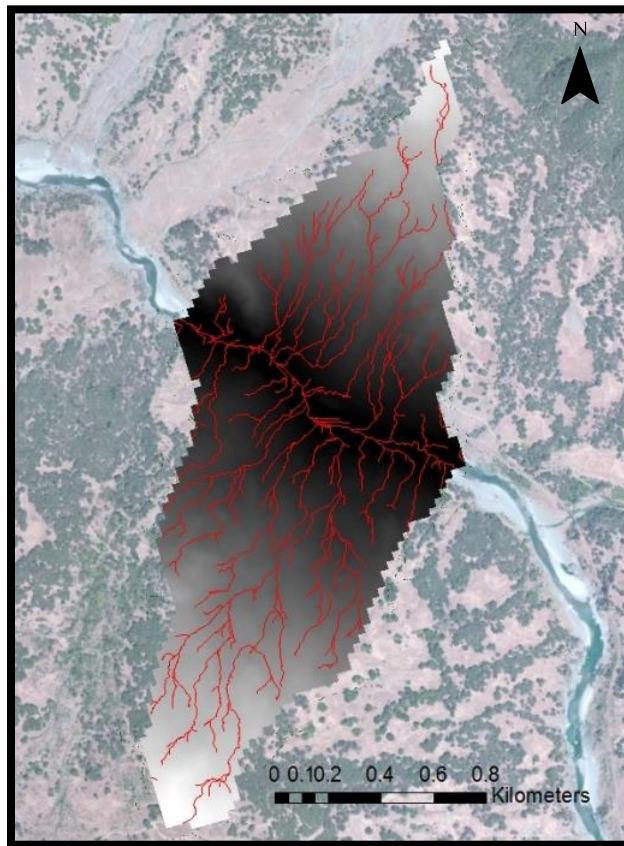


Figure 7. Channel network for Eel River sub-basin extracted in ArcMap.

### ii. Le Seuer sub-basin

The following images depict the raster values of the processing results for the Le Seuer sub-basin, including: slope, curvature, and accumulation area.



Figure 8. From left to right: slope, curvature, and accumulation area rasters for Le Seuer sub-basin.

The quantile-quantile plot below indicates that channelized pixels occur at values of curvature slightly greater than zero.

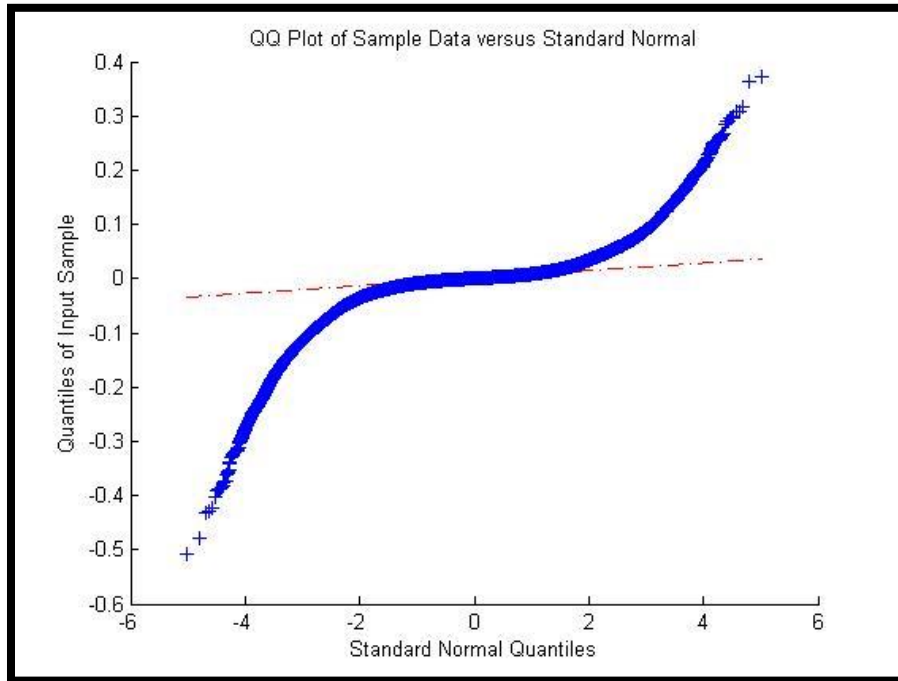


Figure 9. Quantile-quantile plot of curvature values in Le Seuer sub-basin.

Again, the inflection point on the following log-slope versus log-area graph is not very distinct, but there appears to be a small inflection around the 3 value. This corresponds to 1000 area units in the raster, or 9000 sq. meters, as this data is given in 3 meter resolution.

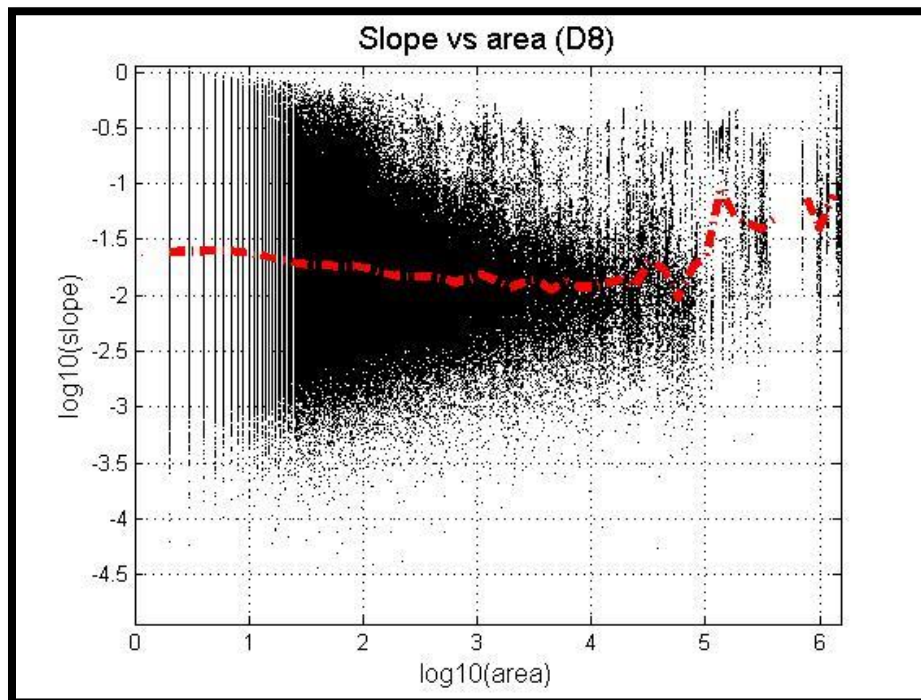


Figure 10. Plot of log-slope versus log-accumulation area. Inflection point is identified as corresponding to the accumulation area equal to 1000 area units (9000 sq. meters for 3-m resolution DEM).



The extracted channel network based on the area threshold is shown below.

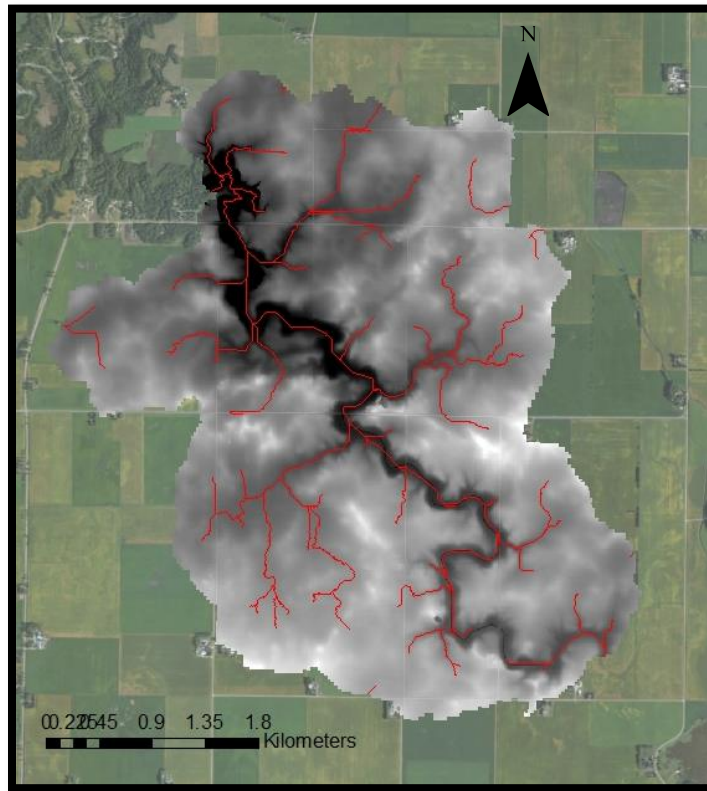


Figure 11. Channel network for Le Seuer sub-basin extracted in ArcMap.

**b. GeoNet processing**

**i. Eel River sub-basin**

The GeoNet tool produces the same information as described above in the section describing the ArcMap results. The outputs from the program are shown below.

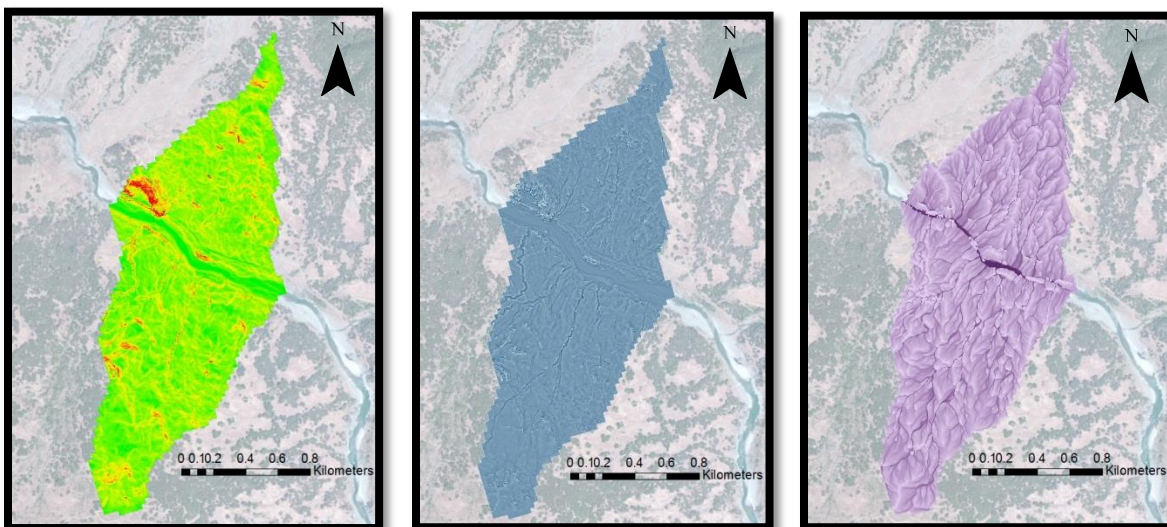


Figure 12. From left to right: slope, curvature, and accumulation area rasters for Eel River sub-basin.

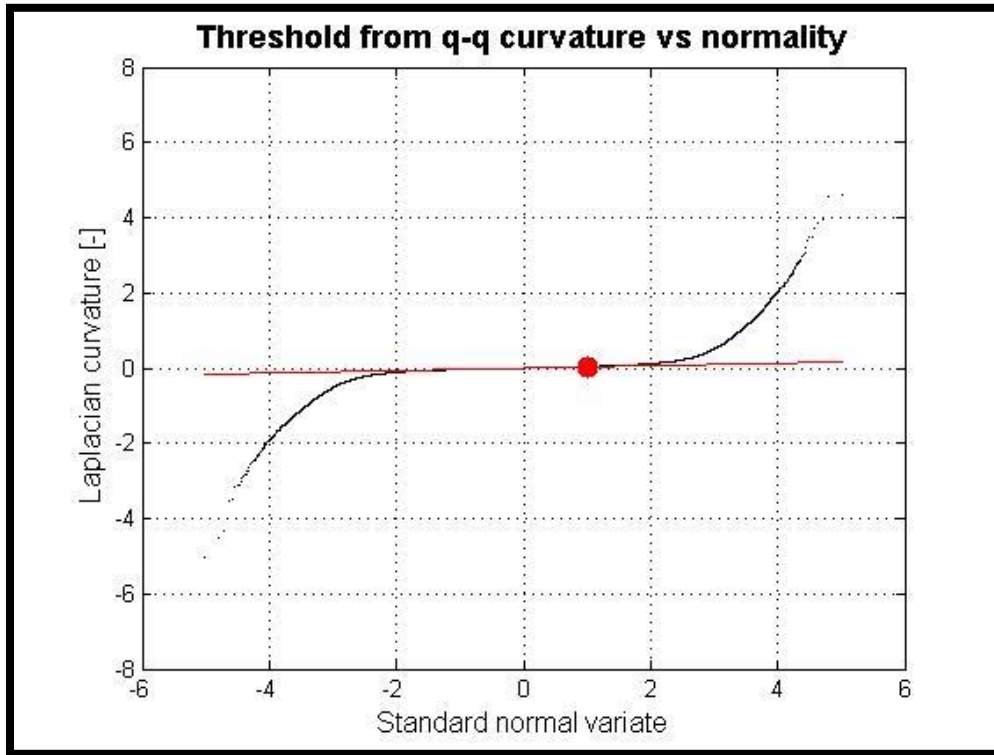


Figure 13. Quantile-quantile plot of curvature values in Eel River sub-basin.

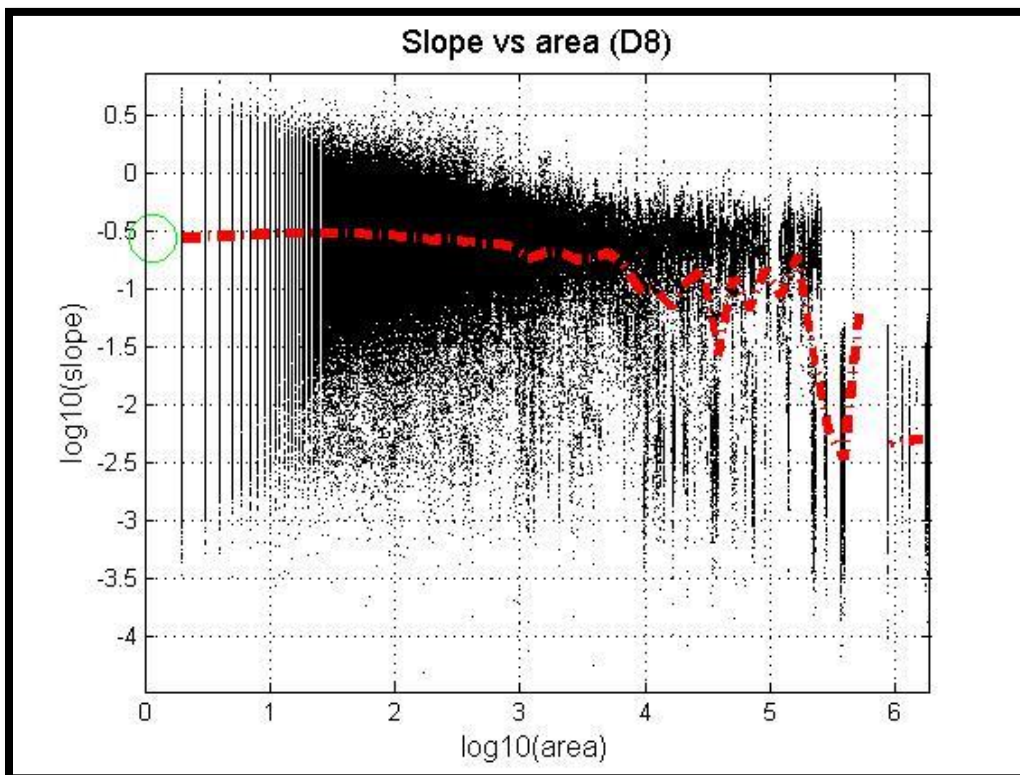


Figure 14. Plot of log-slope versus log-accumulation area.

The inflection point is identified within the GeoNet framework to produce the channel network.

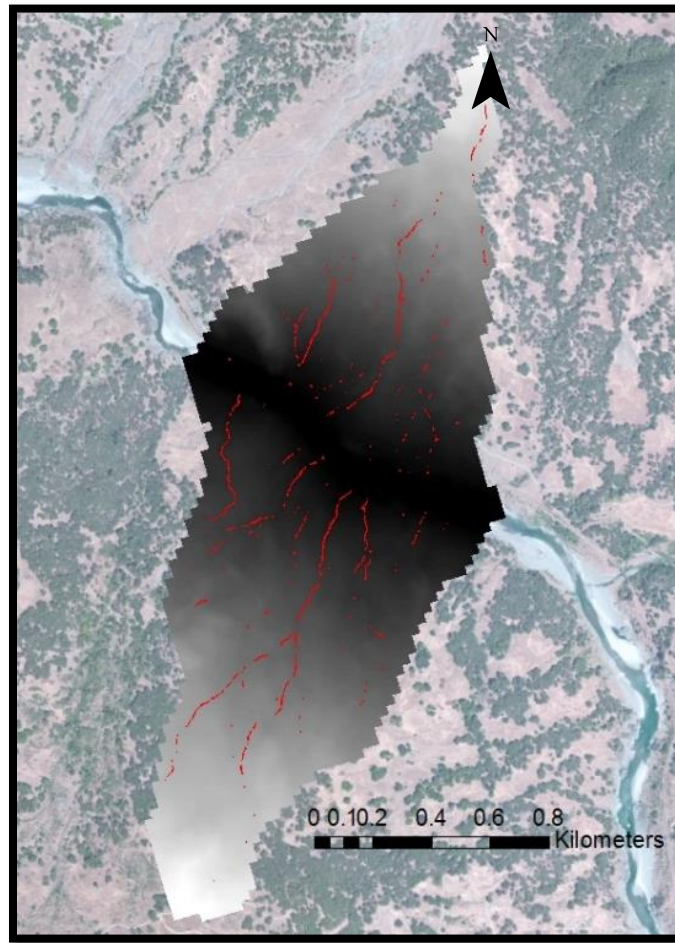


Figure 15. Channel network for Eel River sub-basin extracted by GeoNet.

### ii. Le Seuer sub-basin

Again, the outputs of the GeoNet framework are shown below.

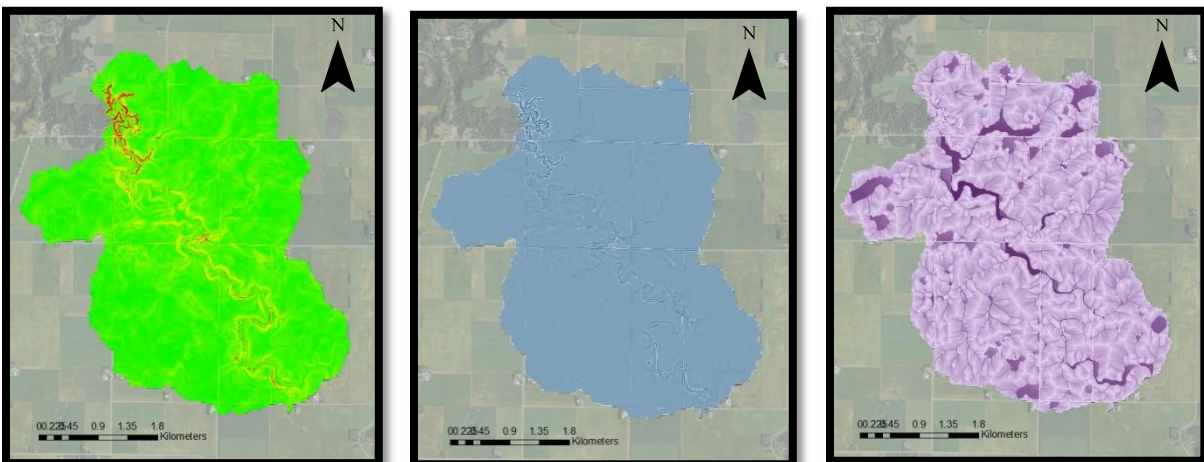


Figure 16. From left to right: slope, curvature, and accumulation area rasters for Le Seuer sub-basin.



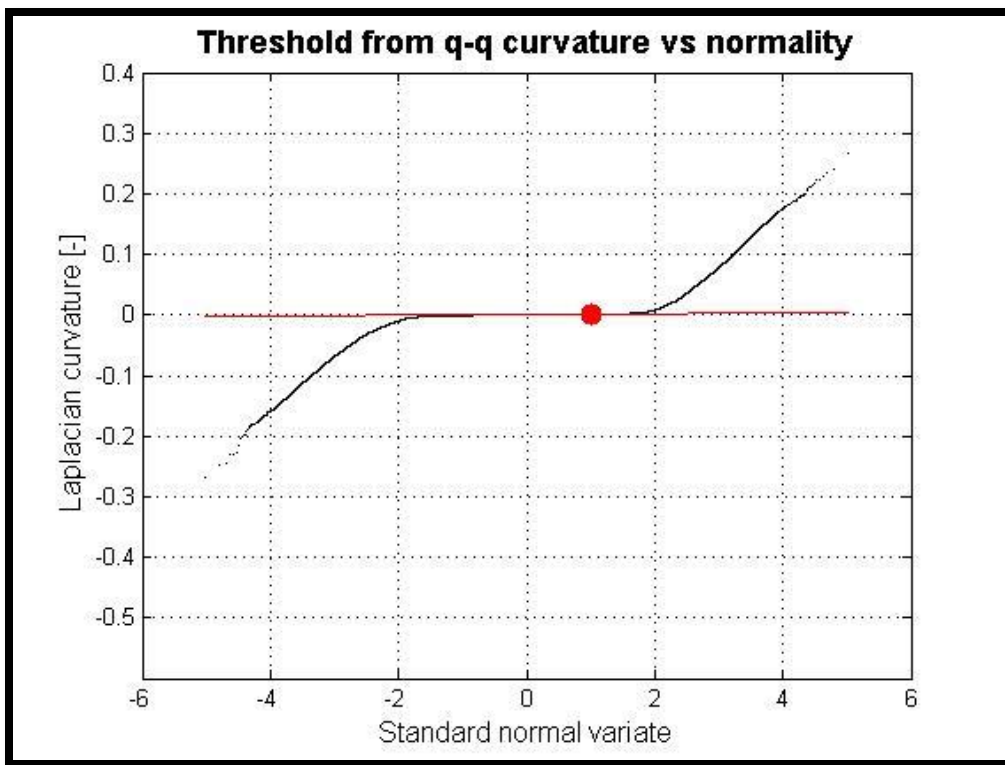


Figure 17. Quantile-quantile plot of curvature values in Le Seuer sub-basin.

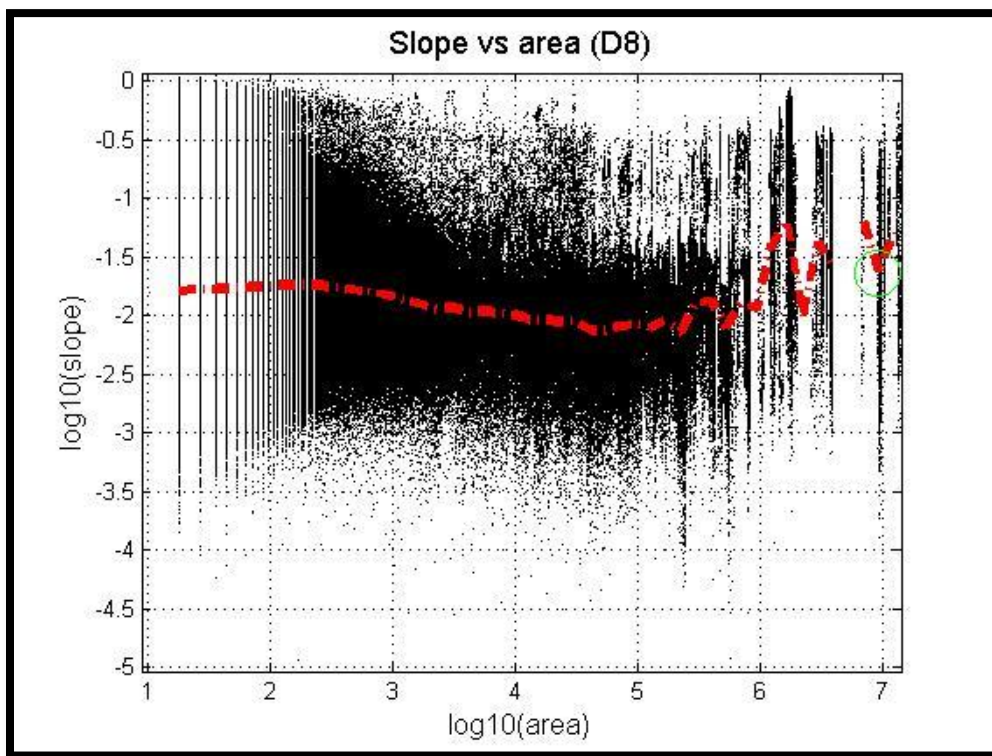


Figure 18. Plot of log-slope versus log-accumulation area.

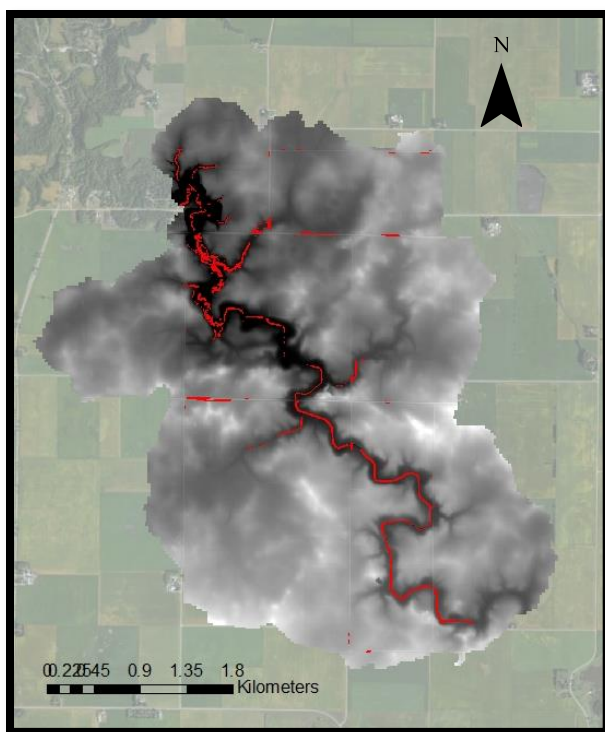


Figure 19. Channel network for Le Seuer sub-basin extracted in GeoNet.

## 5. Discussion of results

### a. Eel River sub-basin

The network results from the different frameworks are starkly different.

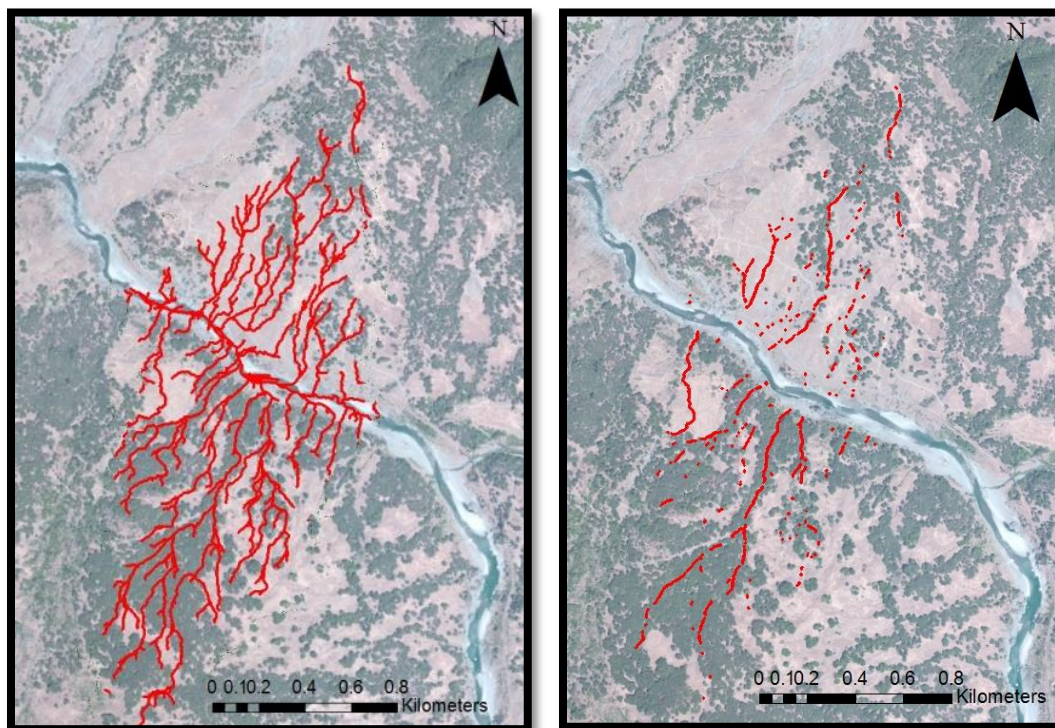


Figure 20. Channel networks for Eel River basin. (a) The network generated by ArcMap is on the left and (b) the network generated by GeoNet is on the right.



From these images, it is apparent there are issues with both of the channel extraction frameworks. The results from ArcMap seem to inflate the channel network density in comparison with the satellite base map imagery, indicating the area threshold chosen based on the log-slope versus log-area graph is too low. The results from GeoNet, in contrast, appear too sparse. In fact, the center channel which is visible from the satellite imagery is not identified as a channel.

For the ArcMap results, if the threshold is increased to 8,000 and 11,000 sq. meters, the channel networks compare to the original (3,000 sq.meters) as follows:

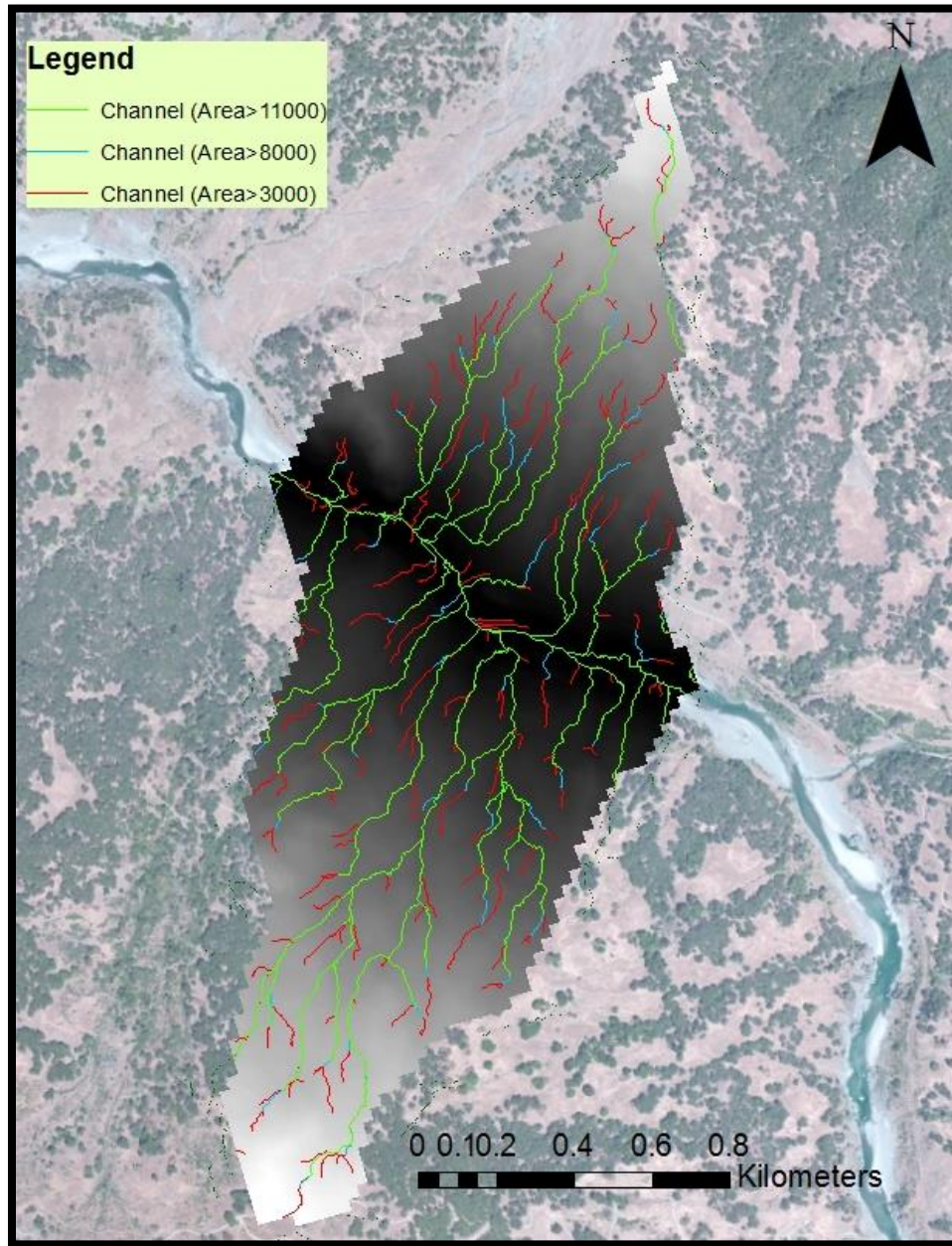


Figure 21. Comparison of channel networks in Eel River sub-basin extracted with ArcMap based on different accumulation area thresholds.

Increasing the threshold area value decreases the density of the channel network slightly. However, in comparison with the satellite imagery, this network still seems too dense, as shown below in Figure 22.

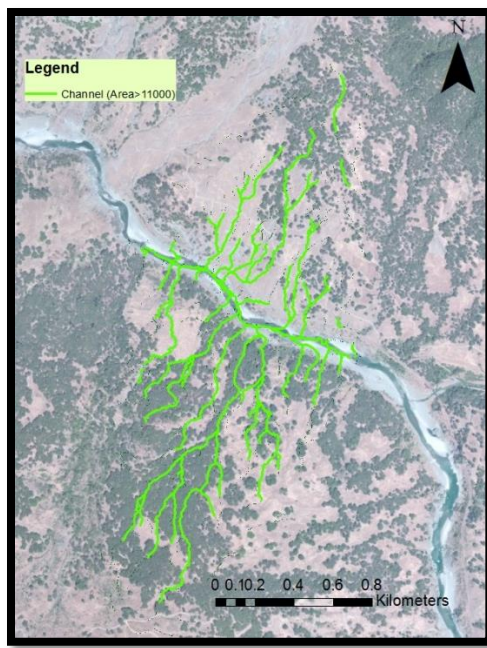


Figure 22. Channel network in Eel River sub-basin extracted in ArcMap at area threshold of 11,000 sq. meters.

**b. Le Seuer sub-basin**

Again, the network results from ArcMap and GeoNet are distinct from each other.

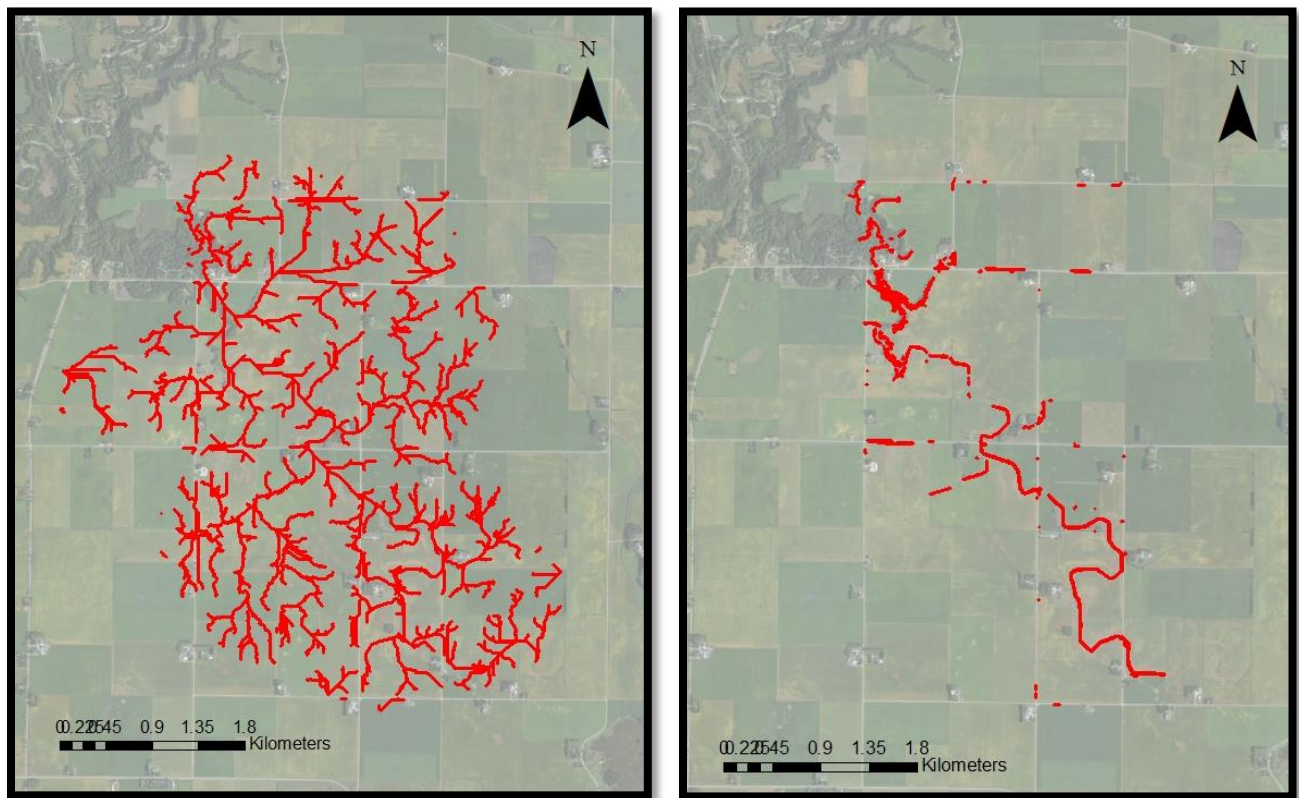


Figure 23. Channel networks for Le Seuer sub-basin. (a)The network generated by ArcMap is on the left and (b) the network generated by GeoNet is on the right.



As in the Eel River channel network, the ArcMap framework seems to overestimate the density of the network. The results from the GeoNet framework appear more accurate, especially overlaid with the digital elevation model as shown in Figure 16. It is likely that these differences result from the differing filtering methods used in ArcMap and GeoNet.

## 6. Conclusions

For the Eel River sub-basin, this analysis shows that neither framework accurately depicts the channel network. However, it may be necessary to have field data of hydrologic channels with which to compare these networks. The dense vegetation present in the satellite imagery available in ArcGIS obscures much of the terrain.

For the Le Seuer sub-basin, both frameworks extracted the major meandering channel that cuts the center of the basin. The GeoNet framework results seem to depict a more accurate overall portrayal of the channel network. The results from both frameworks include roads extracted as channel features as shown in Figure 23 and Figure 24 below. However, the inaccuracy of this is debatable. It may be valid to consider road features as channels, because during precipitation events it is likely that they act as pathways conveying water.



Figure 24. Roads extracted as channel features in Le Seuer sub-basin by (a) ArcMap and (b) GeoNet.

## 7. Future work

This research has been an introduction to my M.S. research. The objectives for my M.S. include a continued study of feature extraction tools, and recommendations for improvements to these tools. Therefore, I will continue to compare and examine filtering methodologies for LIDAR data, as well as feature extraction methods. I endeavor to develop recommendations for different methods based on landscape characteristics and basin size.

## 8. References

- [1] Passalacqua, P., T. Do Trung, E. Foufoula-Georgiou, G. Sapiro, and W. E. Dietrich (2010a), A geometric framework for channel network extraction from lidar: Nonlinear diffusion and geodesic paths, *J. Geophys. Res.*, 115, F01002, doi:10.1029/2009JF001254.
- [2] Roering, J.J., Mackey, B.H., Marshall, J.A., Sweeney, K.E., Deligne, N.I., Booth, A.M., Handwerker, A.L., Cerovski-Darriau C. (2013). You are HERE: connecting the dots with airborne lidar for geomorphic fieldwork. *Geomorphology*, Volume 200, <http://dx.doi.org/10.1016/j.geomorph.2013.04.009>
- [3] Perona, P., and J. Malik (1990), Small-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Mach. Intel.*, 12(7), 629-639.