CSDMS TauDEM Clinic "Hands On" Exercise

David Tarboton, dtarb@usu.edu
Utah State University

Goal

Be able to use TauDEM tools to derive hydrologically useful information from Digital Elevation Models (DEMs)

Purpose

The purpose of this exercise is to introduce Hydrologic Terrain Analysis using TauDEM and to guide you through the steps of running some of the more important functions required to delineate a stream network. Comprehensive documentation on the use of each TauDEM function is given in the online help that is part of the program.

TauDEM (Terrain Analysis Using Digital Elevation Models) is a set of Digital Elevation Model (DEM) tools for the extraction and analysis of hydrologic information from topography as represented by a DEM. This is software developed at Utah State University (USU) for hydrologic digital elevation model analysis and watershed delineation and may be obtained from http://hydrology.usu.edu/taudem/taudem5.0/.

In this exercise, you will perform the following tasks:

- Basic Grid Analysis using TauDEM functions, including.
    - Pit Remove
    - D8 Flow Directions
    - D8 Contributing Area
    - Grid Network
    - D-Infinity flow direction
    - D-infinity Contributing Area
- Stream Network Analysis using TauDEM functions, including
    - Stream Definition by threshold
    - Move Outlets to Streams
    - Stream Reach and Watershed
    - Peuker Douglas
    - Peuker Douglas Stream Definition
- Specialized Grid Analysis using TauDEM functions, including
    - Wetness index derived from slope to area ratio
    - D-Infinity Distance Down

The Logan River watershed is used as an example.

Computer Setup

To complete this exercise, you will need to use the TauDEM 5.0 software as well as MPICH2 software from http://www.mcs.anl.gov/research/projects/mpich2/. You will also need to use visualization software such as ArcGIS or R. Refer to CSDMS Workshop setup instructions at http://hydrology.usu.edu/taudem/taudem5.0/documentation.html for computer setup details.

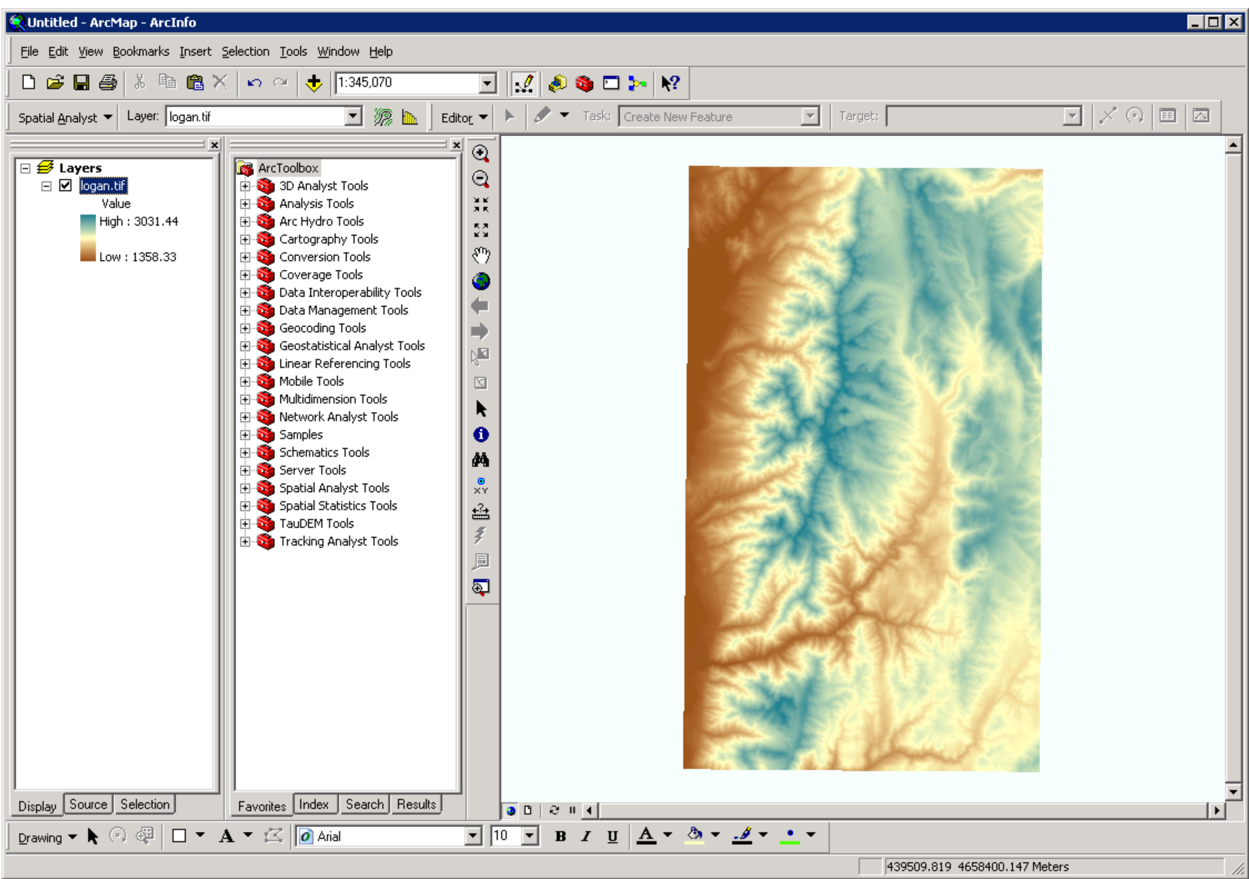
Basic Grid Analysis using TauDEM functions

In this section we illustrate the TauDEM basic grid analysis functions.

1. Download the Logan River example data zip file from http://hydrology.usu.edu/taudem/taudem5.0/LoganDemo.zip. Extract all files from the zip file. Look at the data.

This data was obtained originally from the National Elevation Dataset seamless data server. See appendix 1 for how to obtain US DEM data from the USGS seamless data server and project it to a spatial reference system for the area of interest. Projected data should be used when working with TauDEM because TauDEM uses grid dimensions (cell size) in its length and slope calculations and these will be incorrect if they are not consistent in E-W and N-S directions and in the same units as the vertical units of the DEM.

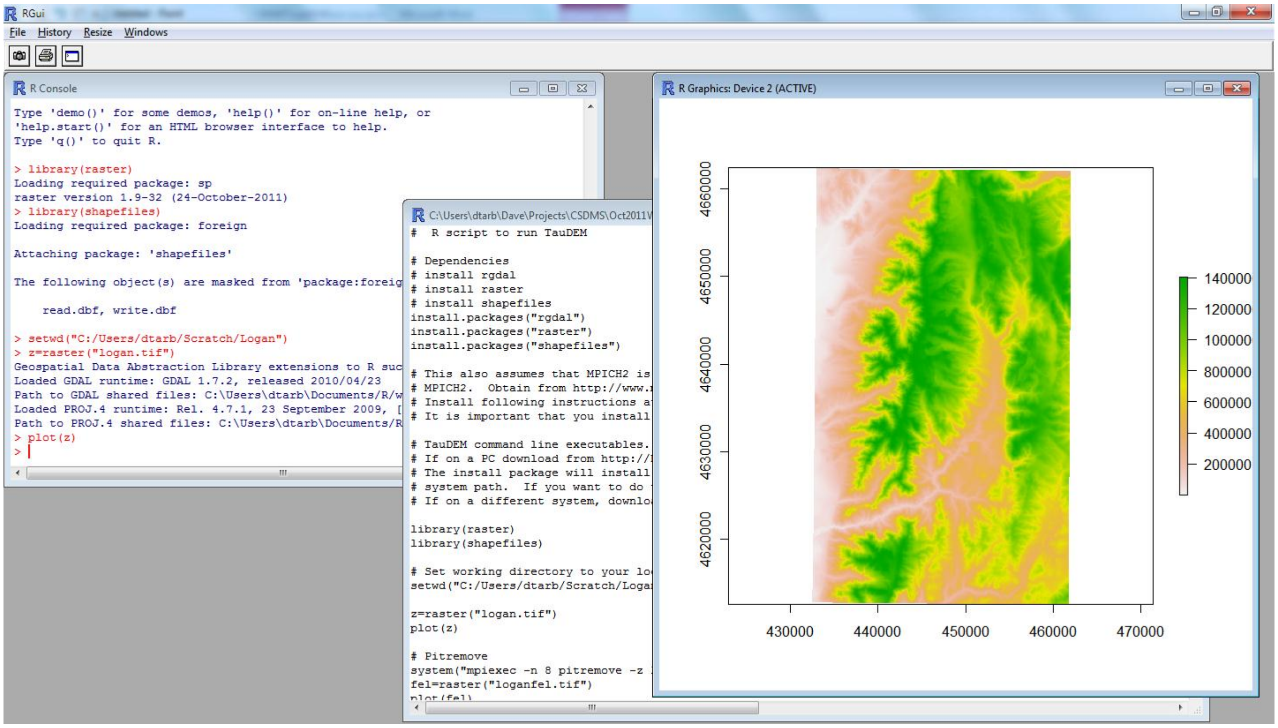
ArcMAP. Add data and adjust the symbology.



R

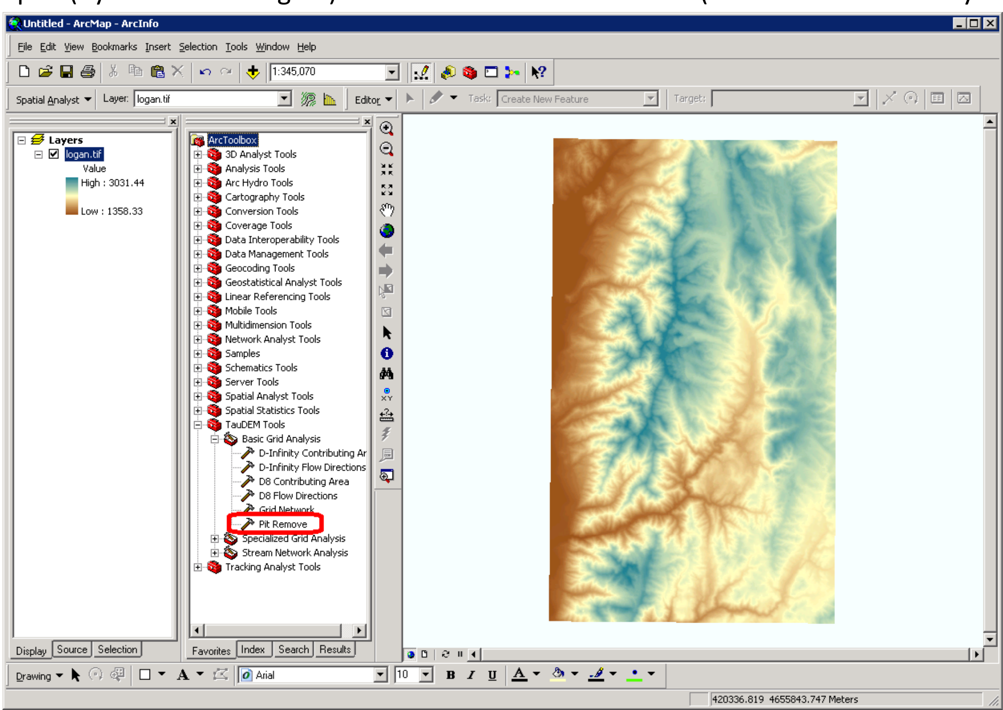
```
z=raster("logan.tif")
```

```
plot(z)
```

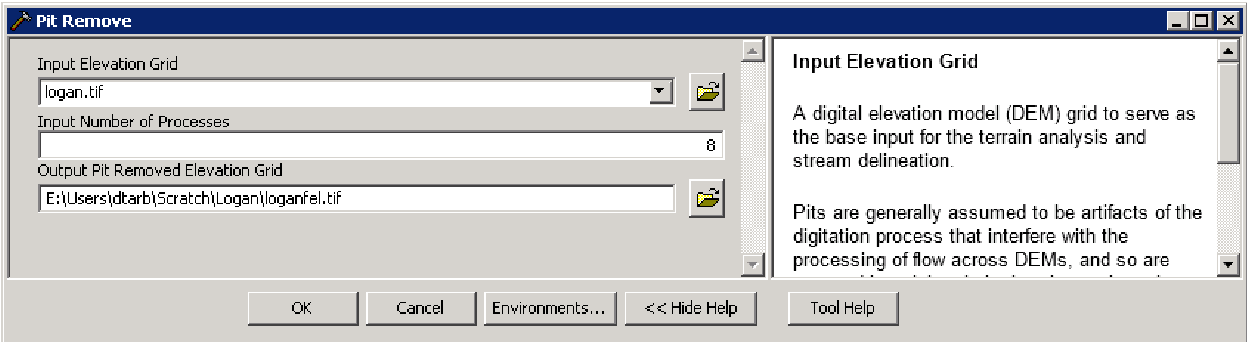


2. The first TauDEM function used is **Pit Remove**. Pits are grid cells surrounded by higher terrain that do not drain. Pit Remove creates a hydrologically correct DEM by raising the elevation of pits to the point where they overflow their confining pour point and can drain to the edge of the domain.

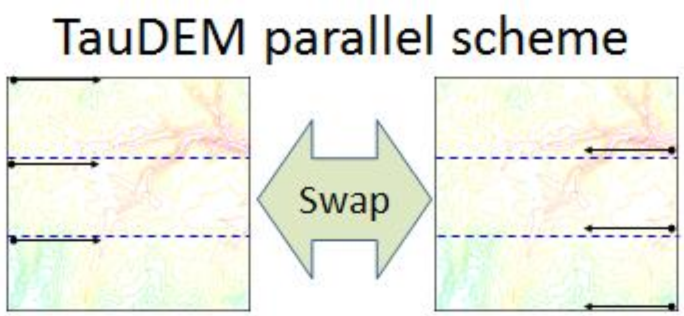
ArcMap. Open (by double clicking on) the TauDEM Pit Remove Tool (in the Basic Grid Analysis set)



Select **logan.tif** for the Input Elevation Grid. Note that the Output Pit Removed Elevation Grid name is automatically filled with loganfel.tif following the file naming convention. Select the Input Number of Processes (I used 8 for a dual quad core PC).

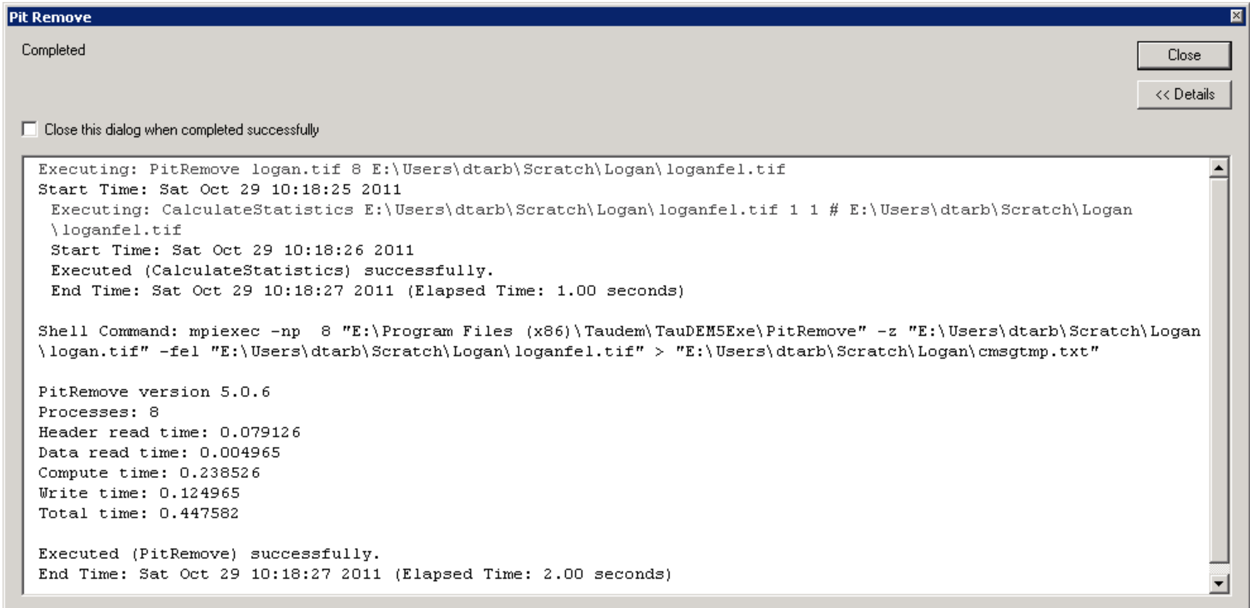


The parallel approach used by TauDEM is illustrated below. The domain is subdivided into row oriented partitions that are each processed independently by separate processes. When the algorithms reach a point where they can proceed no further within the partitions there is a swap step that exchanges information along the boundaries. The algorithms then proceed working within the partitions using new boundary information. This process is iterated until completion. The strategies for sharing information across boundaries and iterating are specific to each algorithm.



The number of processes does not have to be the same as the number of processors on your computer, although generally should be the same order of magnitude. The operating system (and MPICH2) takes care of time sharing between processes, so in cases where some processes are likely to be waiting for other processes to complete there may be a benefit in selecting more processes than physical processors on the computer. However then message passing across the borders is increased. For large datasets, some experimentation as to the number of processes that works best (fastest) is suggested.

Click OK on the Pit Remove tool to run the Pit Remove function for the Logan River DEM. The output dialog reports run statistics that include timing, as well as any error or warning messages.

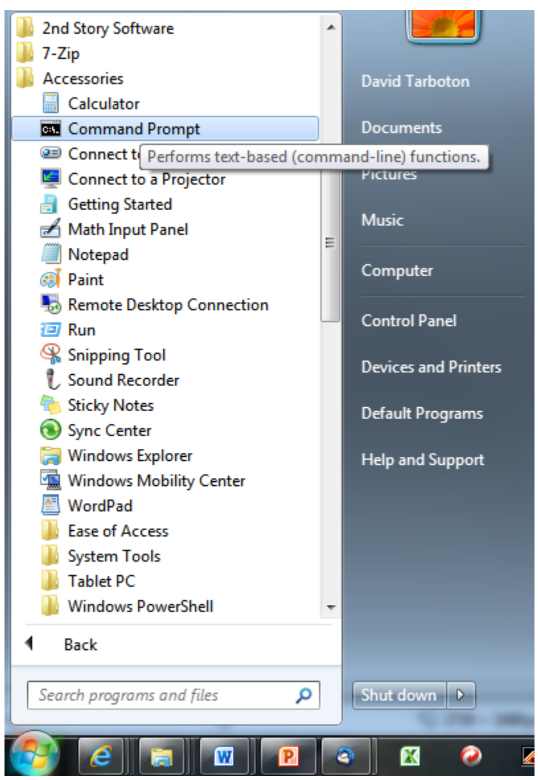


R

```
# Pitremove
system("mpiexec -n 8 pitremove -z logan.tif -fel loganfel.tif")
fel=raster("loganfel.tif")
plot(fel)
```

Command Line

Open a command prompt. Select **Start -> All Programs -> Accessories -> Command Prompt**



In the command prompt type the equivalent on your computer to

```
cd C:\Users\dtarb\Scratch\Demo
```

This changes directory to the folder where you are working

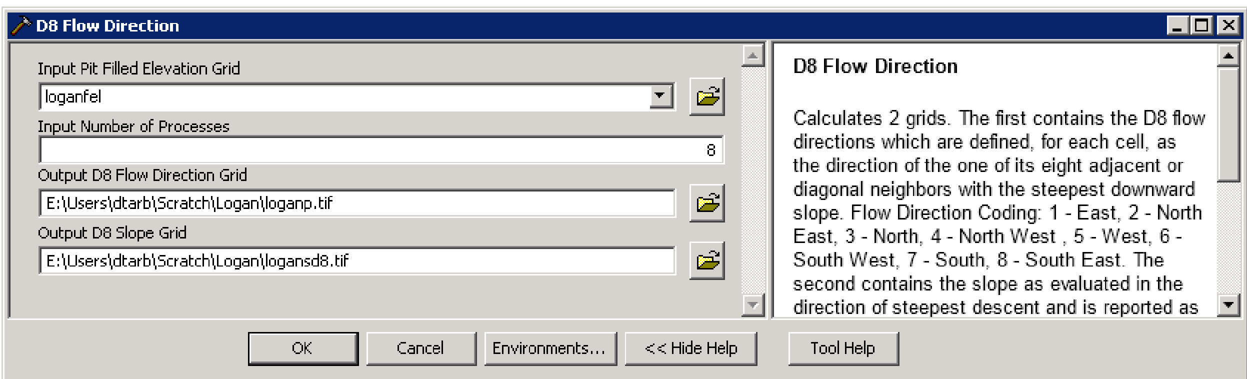
Type (or cut and paste from here) into the command prompt

```
mpiexec -n 8 pitremove -z logan.tif -fel loganfel.tif
```

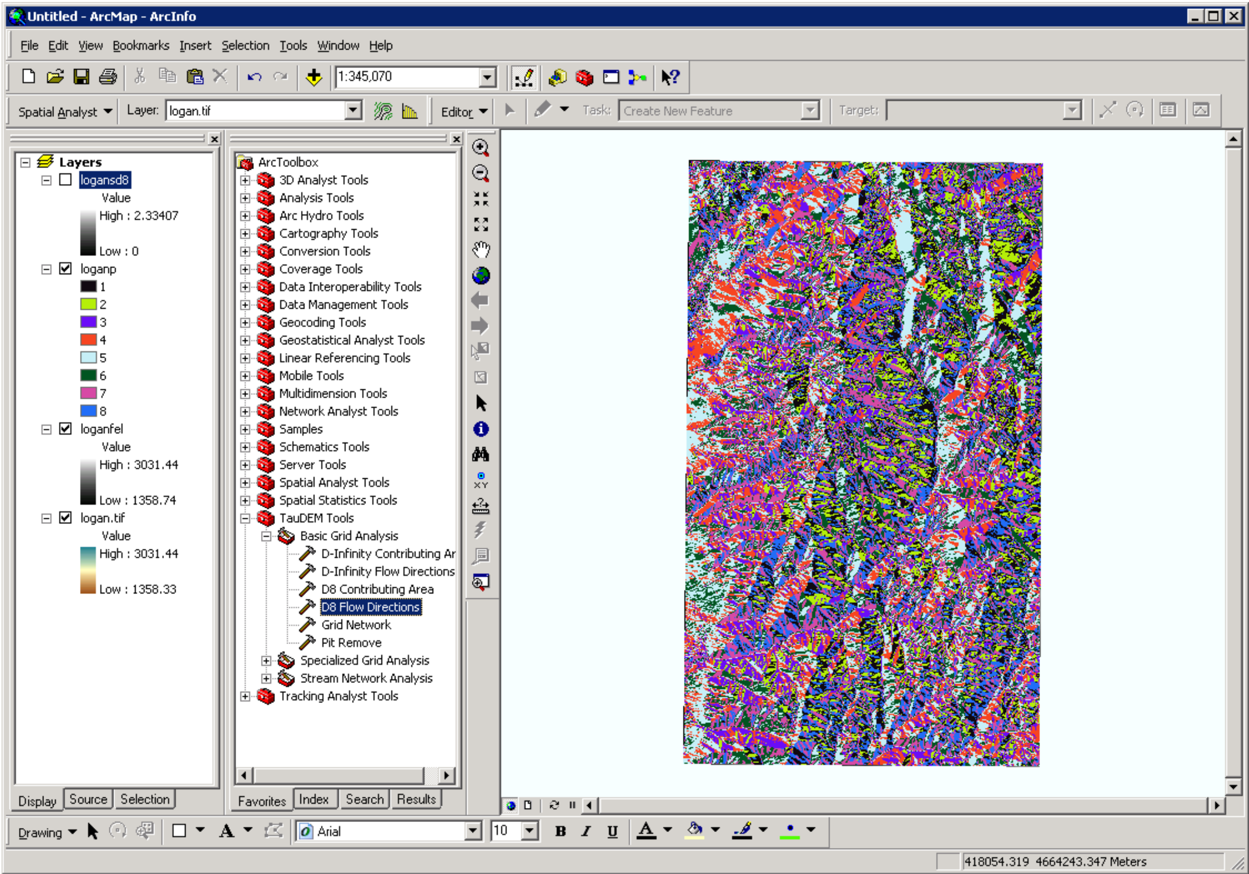
Note. Command line commands, are exactly the same as the string argument in R system calls so are not repeated in what follows. You can learn more about running TauDEM from the command line at: http://hydrology.usu.edu/taudem/taudem5.0/TauDEM5LineGuide.pdf

3. The next function to run is **D8 Flow Direction**. This takes as input the hydrologically correct elevation grid and outputs D8 flow direction and slope for each grid cell.

ArcMap



The resulting D8 flow direction grid (grid has suffix p) is illustrated. This is an encoding of the direction of steepest descent from each grid cell using the numbers 1 to 8 per the embedded help above. This is the simplest model of the direction water would flow over the terrain.



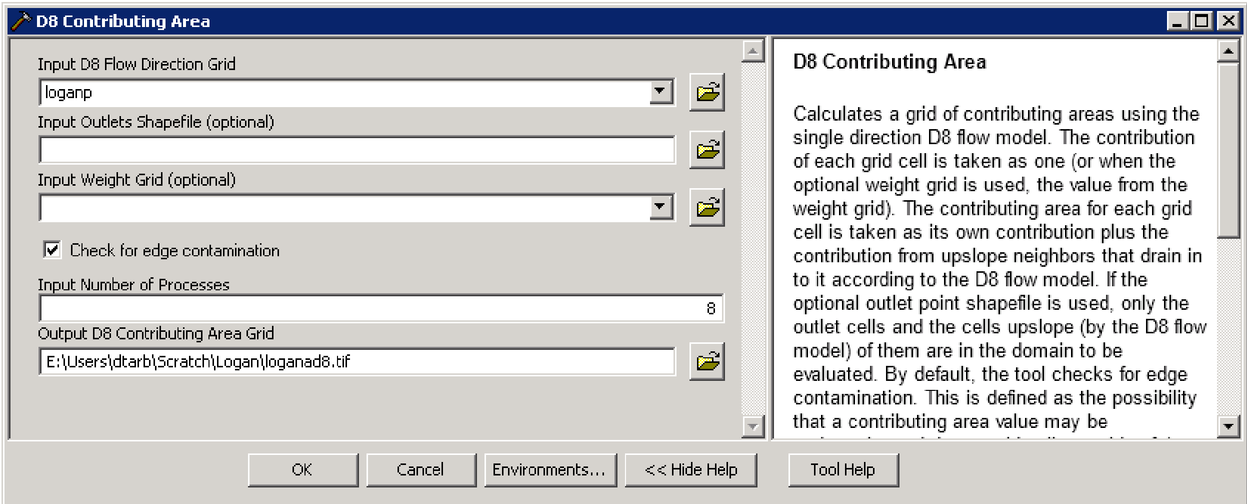
R

```
# D8 flow directions
system("mpiexec -n 8 D8Flowdir -p loganp.tif -sd8 logansd8.tif -fel
loganfel.tif",show.output.on.console=F,invisible=F)
p=raster("loganp.tif")
plot(p)
sd8=raster("logansd8.tif")
plot(sd8)
```

Note that on the R system calls there is the option to include "show.output.on.console=F", and "invisible=F" in the arguments. TauDEM functions write to stderr and stdout. This output gives progress information and an indication of how long each program may take. For long tasks it is helpful to see this to get a sense of progress and get reassurance that the program has not crashed. These arguments enable this, rather than this output being held by R for display only when the task completes. These arguments may be included in any R system call, although this was for expedience not done in many of the calls below.

4. The next function to run is **D8 Contributing Area**. This counts the number of grid cells draining through (out of) each grid cell based on D8 flow directions.

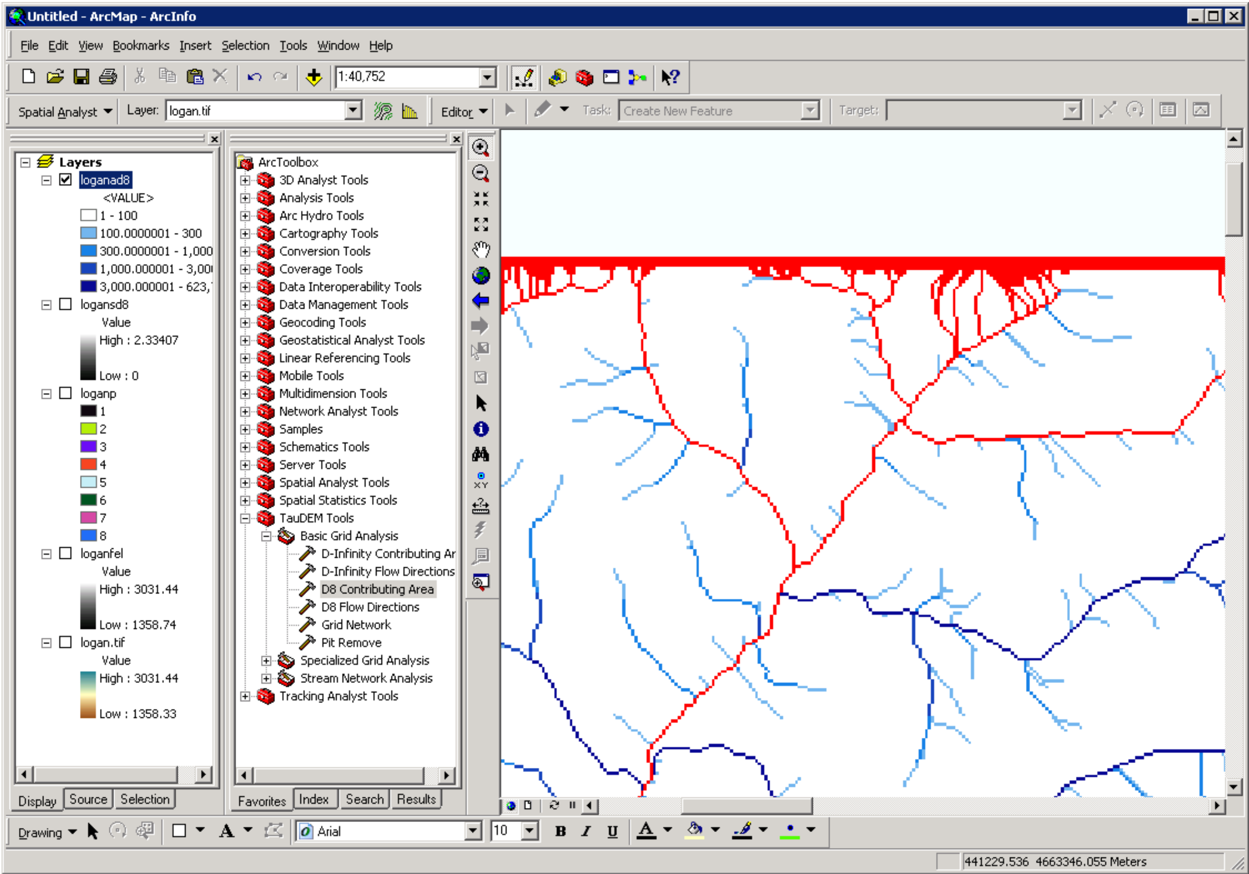
ArcMap



There are options to specify outlets and an input weight grid that were not used here. These are detailed in the tool help and allow calculations to be restricted to the area upstream of designated outlets (specified as a shapefile) and to accumulate an input weight field, rather than just counting contributing area as a number of grid cells.

There is also an option to check for **edge contamination**. Edge contamination is a problem that can occur in the calculation of contributing area when flow is inwards from the boundary of the terrain. The computer does not know what the inflowing contributing area at the edge is, so evaluates the contributing area that may be impacted by this unknown area as no data. The result is streaks that enter the domain along flow paths. This is a desired result as it guards against missing parts of the watershed, but it can and should be turned off if the DEM has been clipped to a watershed boundary. Red has been used to display no data to illustrate edge contamination.

A logarithmic scale is often best to render contributing area values as in the illustration below.

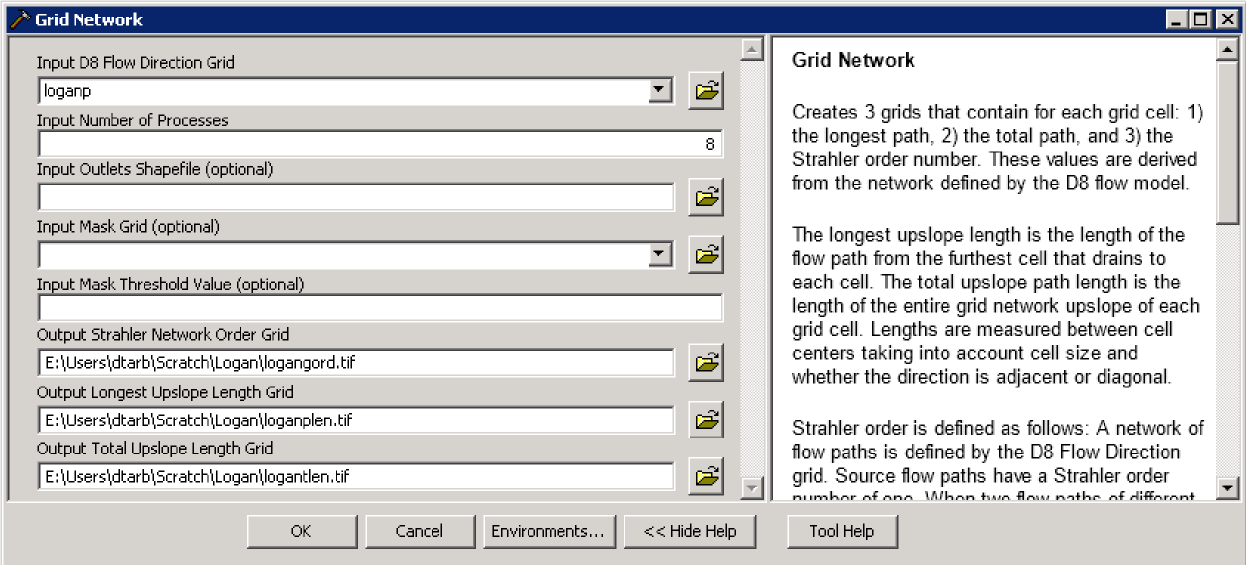


R

```
# Contributing area
system("mpiexec -n 8 AreaD8 -p loganp.tif -ad8 loganad8.tif")
ad8=raster("loganad8.tif")
plot(log(ad8))   # Use log scale for plotting
zoom(log(ad8))
```

5. The **Grid Network** function outputs three grids: (1) the longest flow path along D8 flow directions to each grid cell, (2) the total length of all flow paths that end at each grid cell, and (3) the grid network order. This is obtained by applying the Strahler stream ordering system to the network defined starting at each grid cell.

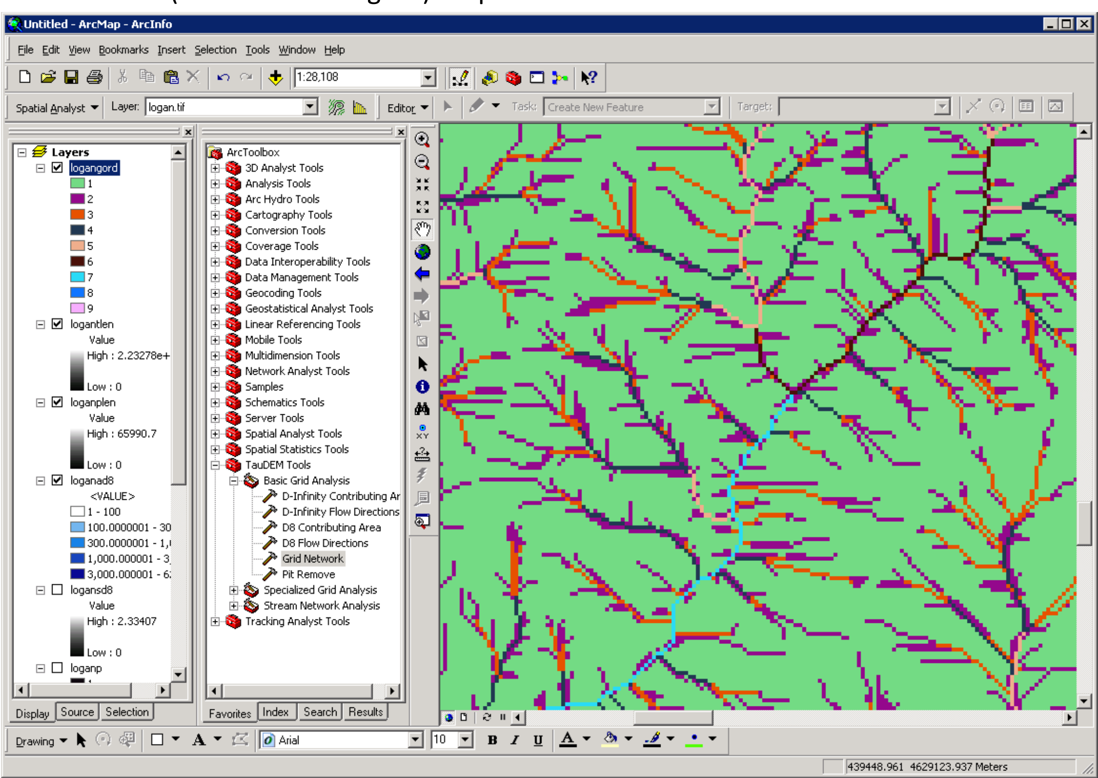
ArcMap



R

```
# Grid Network
system("mpiexec -n 8 Gridnet -p loganp.tif -gord logangord.tif -plen
loganplen.tif -tlen logantlen.tif")
gord=raster("logangord.tif")
plot(gord)
zoom(gord)
```

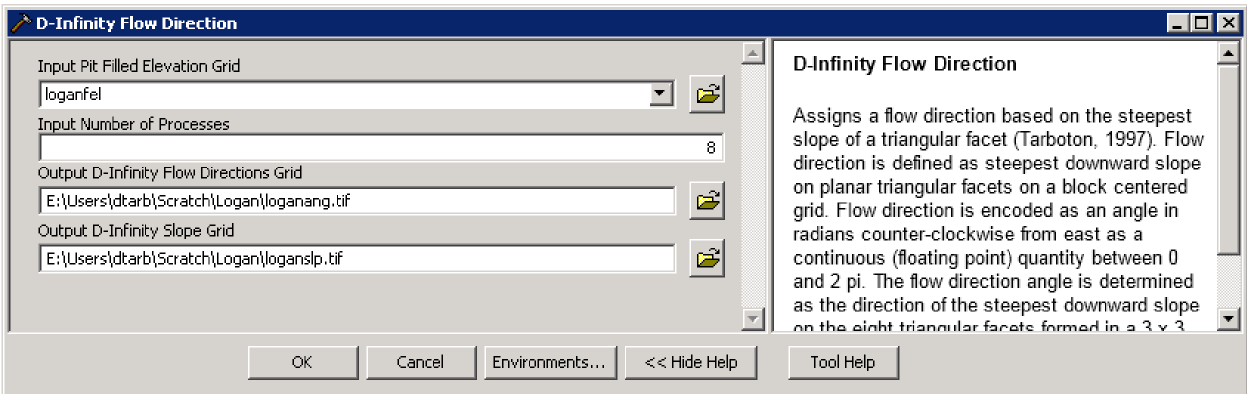
Grid Network Order (file name suffix gord) output from Grid Network is illustrated:



The functions above used the D8 flow model that represents flow from each grid cell to one neighbor. TauDEM also uses the D∞ (D-Infinity) flow model that calculates the steepest outwards flow direction using triangular facets centered on each grid cell and apportions flow between neighboring grid cells based on flow direction angles.

6. The **D-Infinity Flow Direction** function is the starting point for all D-Infinity work. It calculates the D-Infinity flow directions for use in other TauDEM functions requiring D-infinity flow direction input.

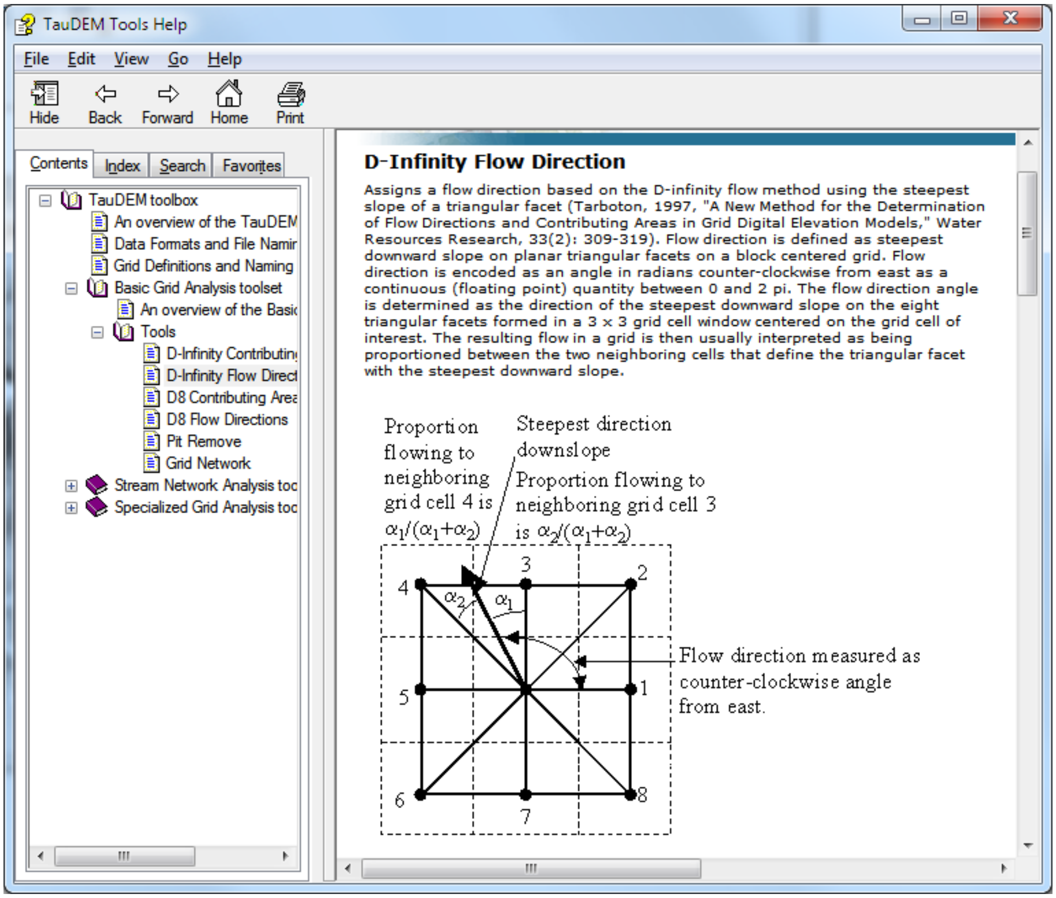
ArcMap



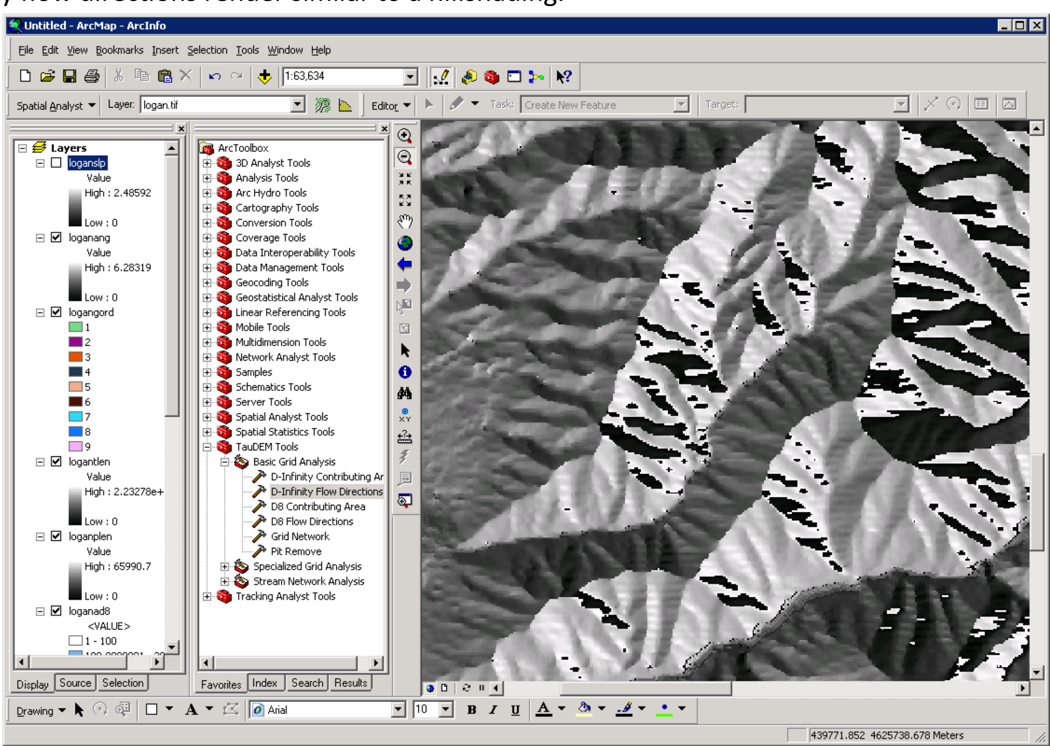
R

```
# DInf flow directions
system("mpiexec -n 8 DinfFlowdir -ang loganang.tif -slp loganslp.tif
-fel loganfel.tif",show.output.on.console=F,invisible=F)
ang=raster("loganang.tif")
plot(ang)
slp=raster("loganslp.tif")
plot(slp)
```

D-Infinity flow directions are encoded as angles counter clockwise from East in Radians as illustrated in the help.

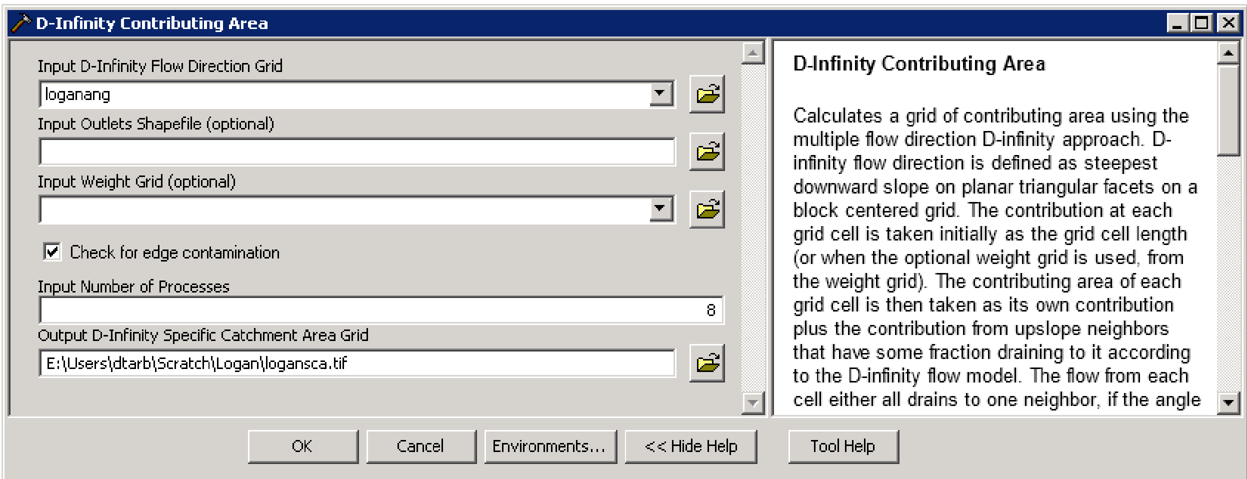


D-Infinity flow directions render similar to a hillshading.



7. **The D-Infinity Contributing Area** function evaluates contributing area using the D-Infinity model based on flow being shared between grid cells proportional to the angle to the steepest downslope direction. This is designed to represent specific catchment area within dispersed flow over a smooth topographic surface.

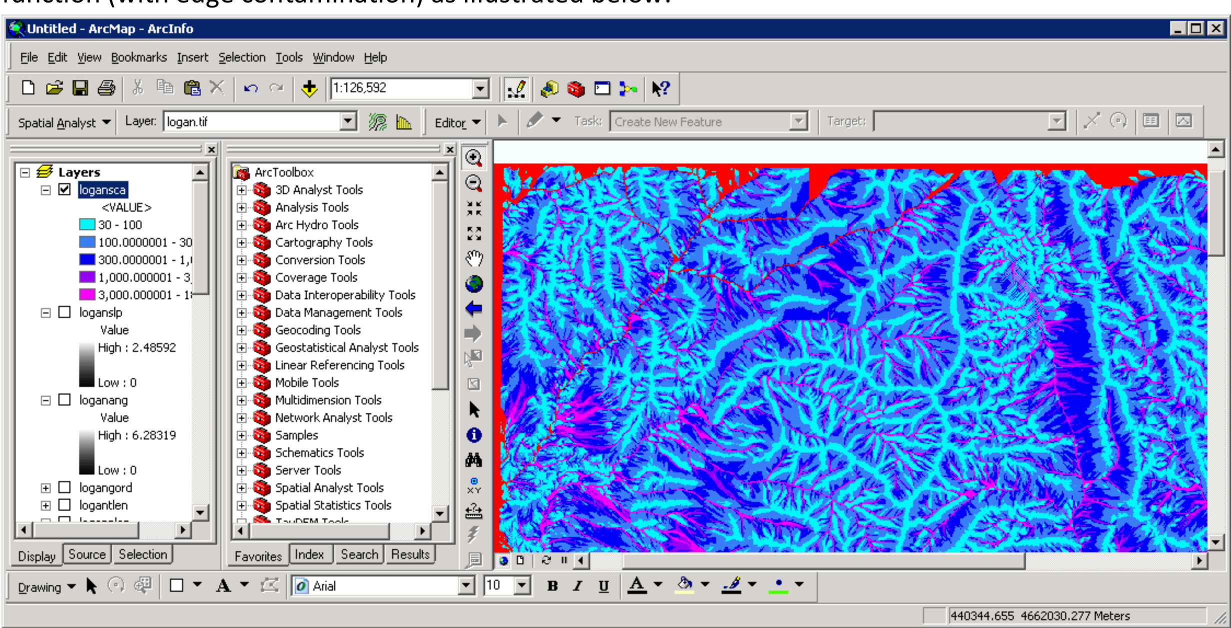
ArcMap



R

```
# Grid Network
system("mpiexec -n 8 AreaDinf -ang loganang.tif -sca logansca.tif")
sca=raster("logansca.tif")
plot(log(sca))
zoom(log(sca))
```

The result from running this is specific catchment area obtained from the D-Infinity contributing area function (with edge contamination) as illustrated below.

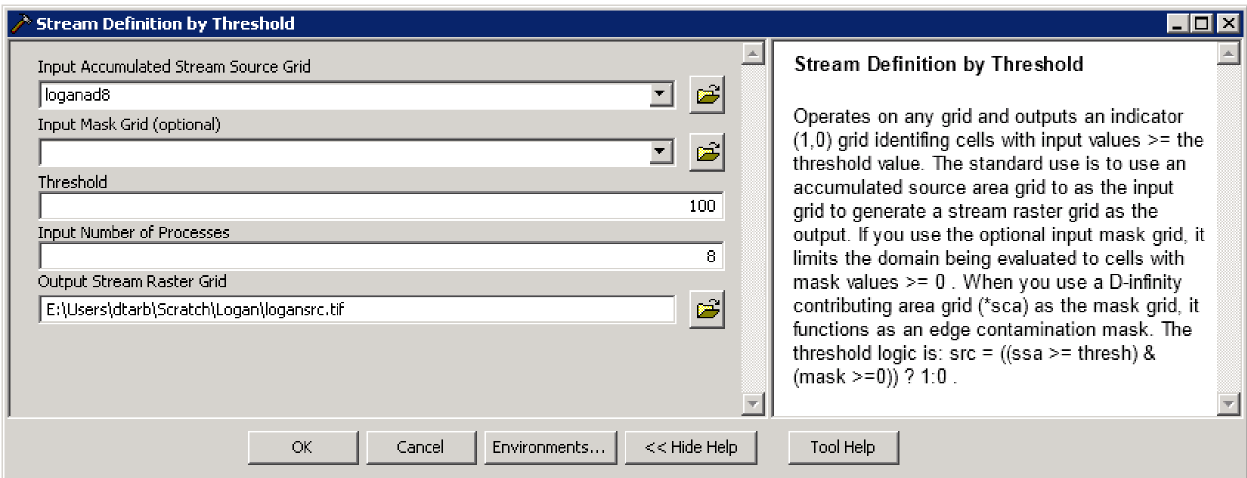


Stream Network Analysis using TauDEM functions

TauDEM provides a number of methods for delineating and analyzing stream networks and watersheds. The simplest stream network delineation method uses a threshold on contributing area.

8. **Stream Definition by Threshold.** This function defines a stream raster grid (src suffix) by applying a threshold to the input. In this case the input is a D8 contributing area grid and a threshold of 100 grid cells has been used.

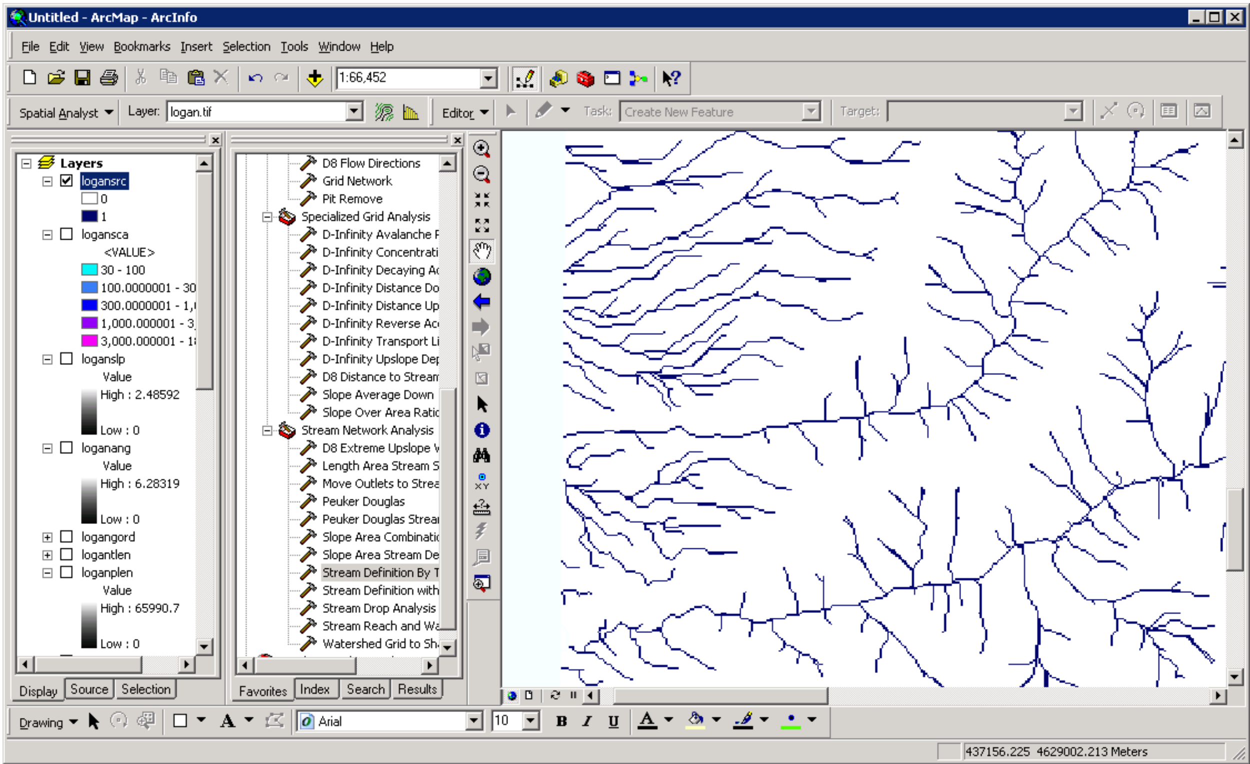
ArcMap



R

```
# Threshold
system("mpiexec -n 8 Threshold -ssa loganad8.tif -src logansrc.tif -
thresh 100")
src=raster("logansrc.tif")
plot(src)
zoom(src)
```

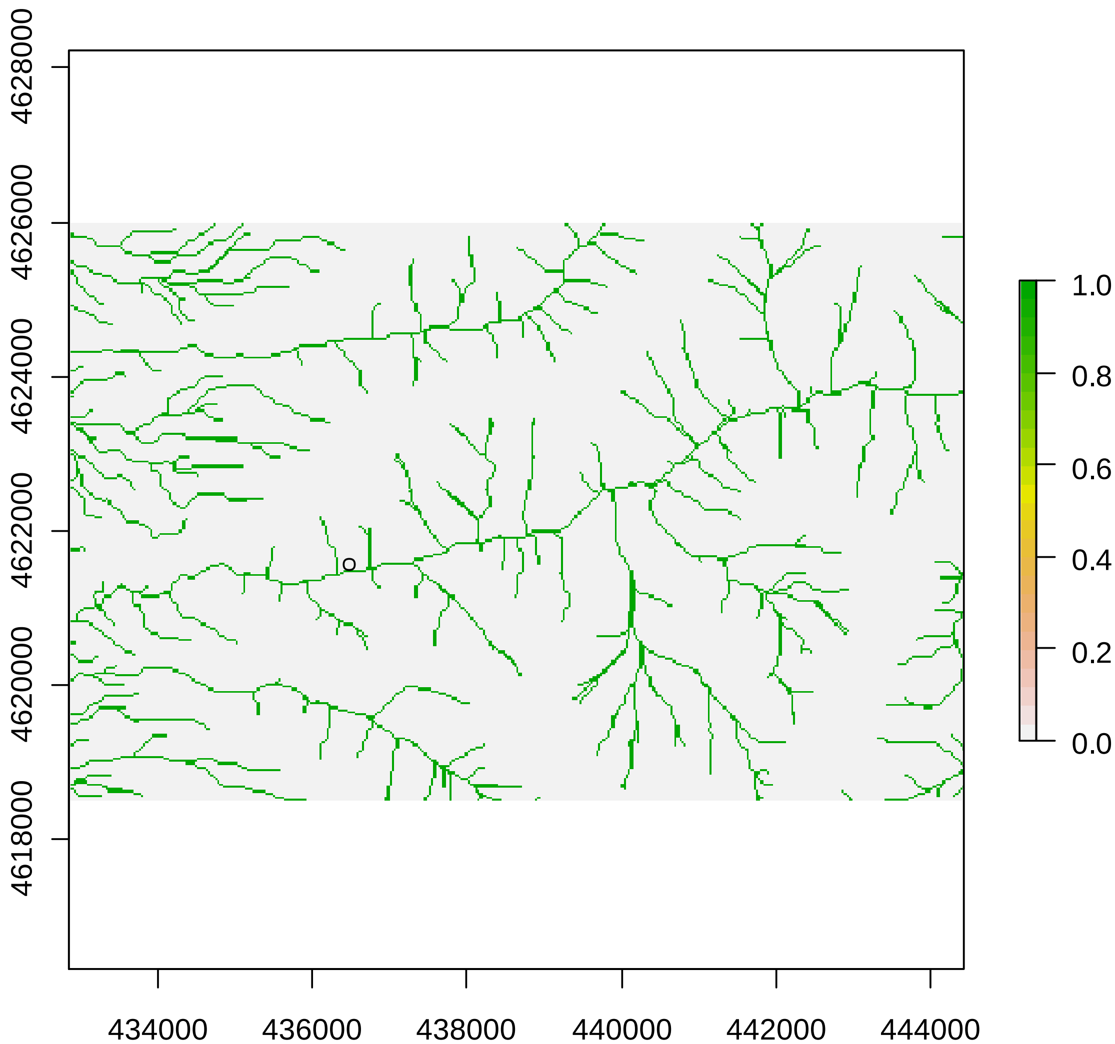
The result depicts the stream network as a binary grid (but is not logically connected as a network shapefile yet).



9. **Identify a watershed outlet**. It is common to want to delineate watersheds upstream of an outlet, say a USGS stream gauge. This requires that gauge locations be precisely located on streams as rendered from the DEM. Due to inaccuracies in gauge locations and DEM stream delineation it is common for gauge locations not to be precisely on streams delineated from the DEM. The move Outlets to streams function slides gauge locations downslope following D8 flow directions until a stream (as defined by a stream raster grid) is encountered. An outlet shapefile can be created by clicking on or near a stream. I did this using the R function below.

R

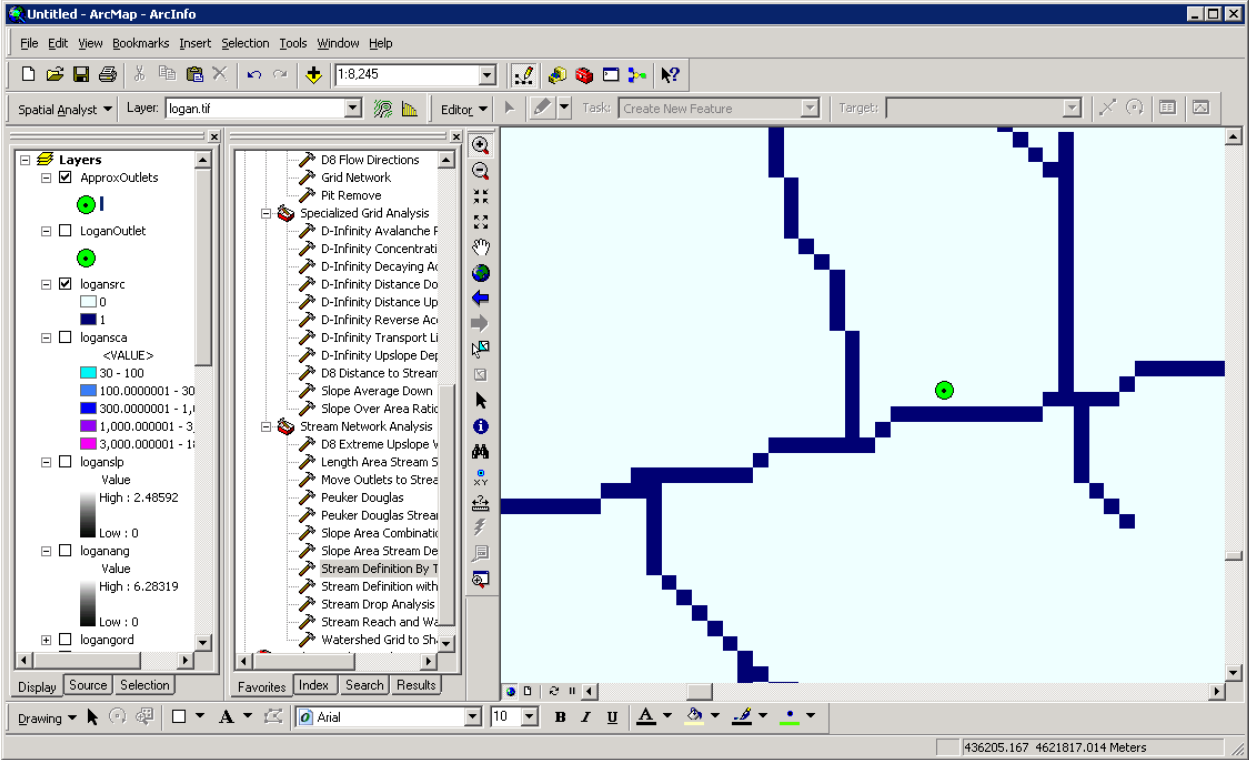
```
# a quick R function to write a shapefile
makeshape.r=function(sname="shape",n=1)
{
xy=locator(n=n)
points(xy)

#Point
dd <- data.frame(Id=1:n,X=xy$x,Y=xy$y)
ddTable <- data.frame(Id=c(1),Name=paste("Outlet",1:n,sep=""))
ddShapefile <- convert.to.shapefile(dd, ddTable, "Id", 1)
write.shapefile(ddShapefile, sname, arcgis=T)
}

makeshape.r("ApproxOutlets")
```



This can also be done in ArcGIS using standard shape editing functionality.

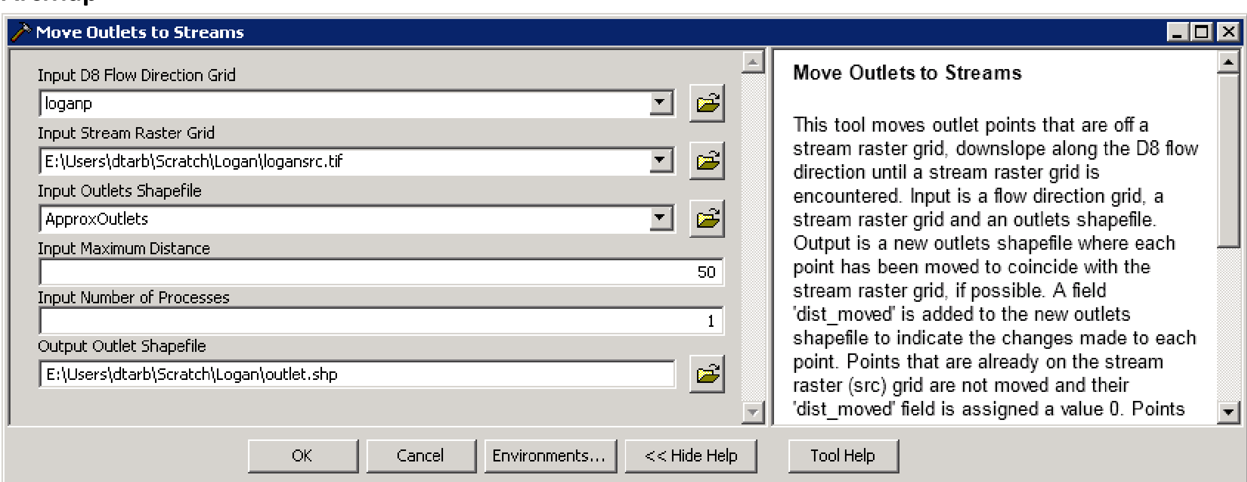
ArcMap

Add the **ApproxOutlets.shp** file to ArcMap and zoom in to the area around it. Change the symbology if necessary.



10. The **Move Outlets to Streams** function is used to move the outlets to the streams.

ArcMap

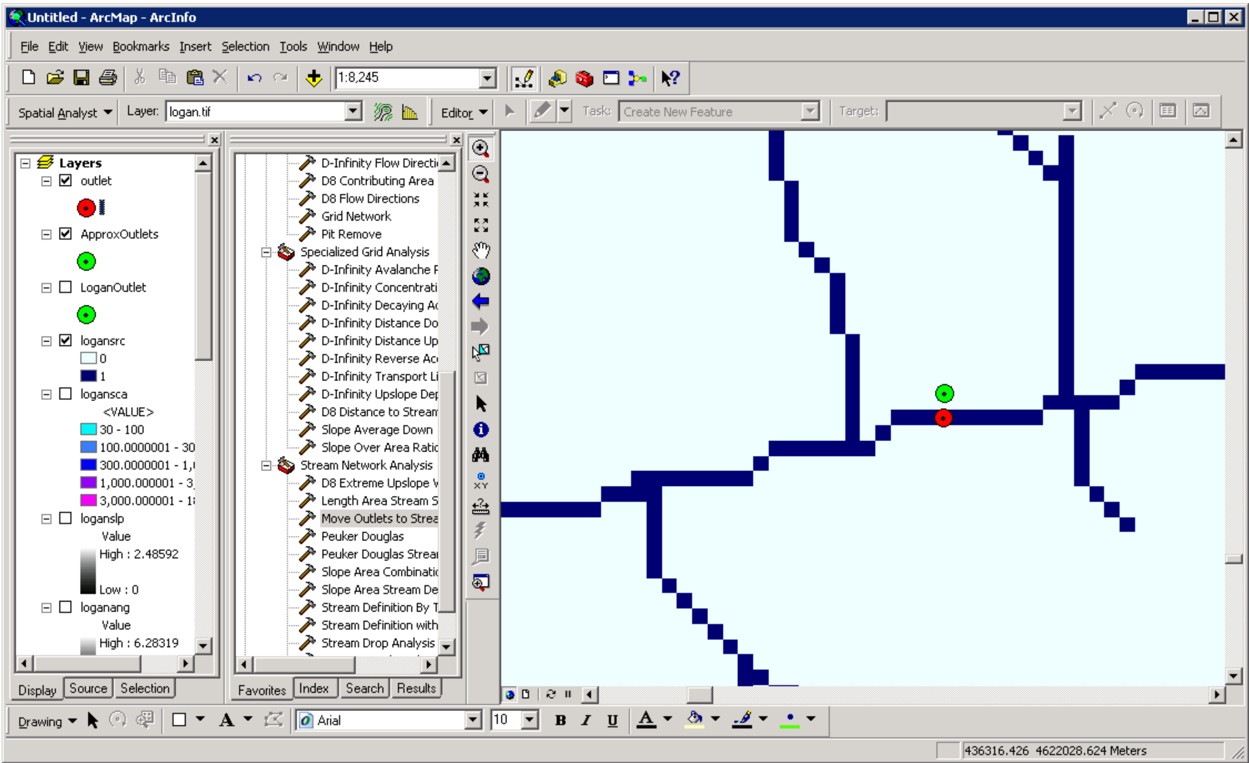


R

```
# Move Outlets
system("mpiexec -n 8 moveoutletstostreams -p loganp.tif -src
logansrc.tif -o approxoutlets.shp -om Outlet.shp")
outpt=read.shp("outlet.shp")
```

```
plot(src)
points(outpt$shp[2],outpt$shp[3],pch=19,col=2)
```

Visualize the **outlet.shp** shapefile. Notice (below) how the outlet has been moved to coincide with the stream.

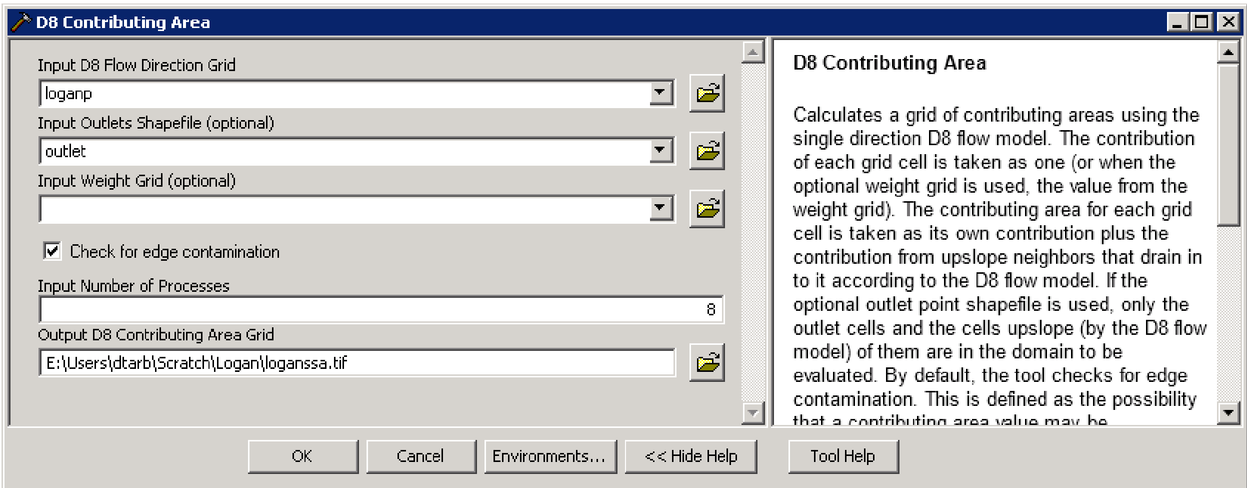


It is somewhat overkill to use 8 processors to move one outlet point, but this is illustrative of how this could be used for many more points. Upon adding the moved outlet you may have received an ArcMAP unknown spatial reference warning. Be aware in using TauDEM that TauDEM does not do any spatial reference (projection) conversions. Therefore all data needs to be in the same spatial reference system. TauDEM does copy the spatial reference information from input grids to output grids, but does not do this for shapefiles.

With the outlet positioned on the stream the stream network upstream of the outlet can be delineated.

11. Once the outlet has been placed exactly on the stream paths, the **D8 Contributing Area** function is run again, but specifying an outlet shapefile to evaluate contributing area and effectively identify the watershed upstream of the outlet point (or points for multiple outlets).

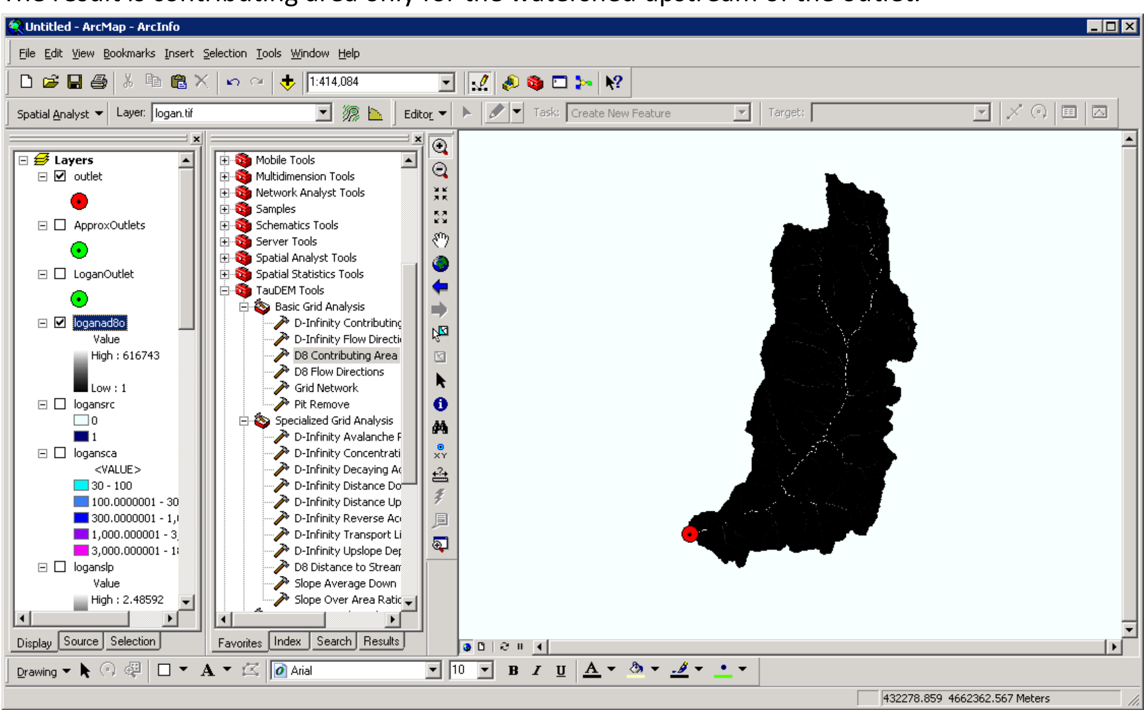
ArcMap



R

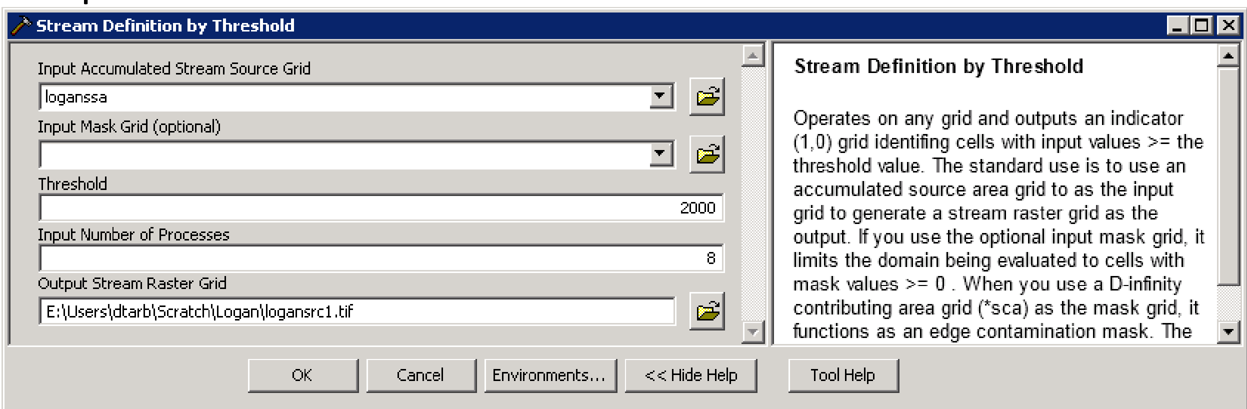
```
# Contributing area upstream of outlet
system("mpiexec -n 8 Aread8 -p loganp.tif -o Outlet.shp -ad8
loganssa.tif")
ssa=raster("loganssa.tif")
plot(ssa)
```

The result is contributing area only for the watershed upstream of the outlet.



12. The next step is to use the **Stream Definition By Threshold** function to define streams using a specified contributing area threshold. Here a threshold of 2000 grid cells is arbitrarily chosen

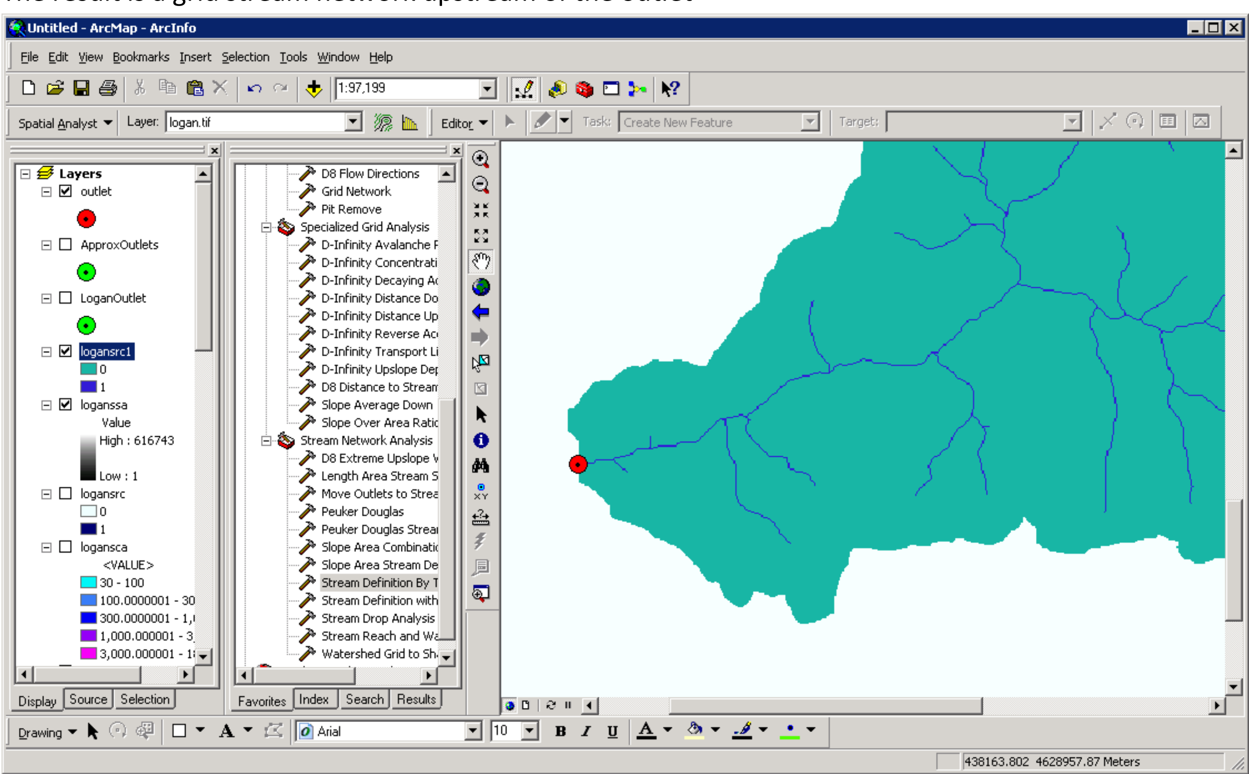
ArcMap



R

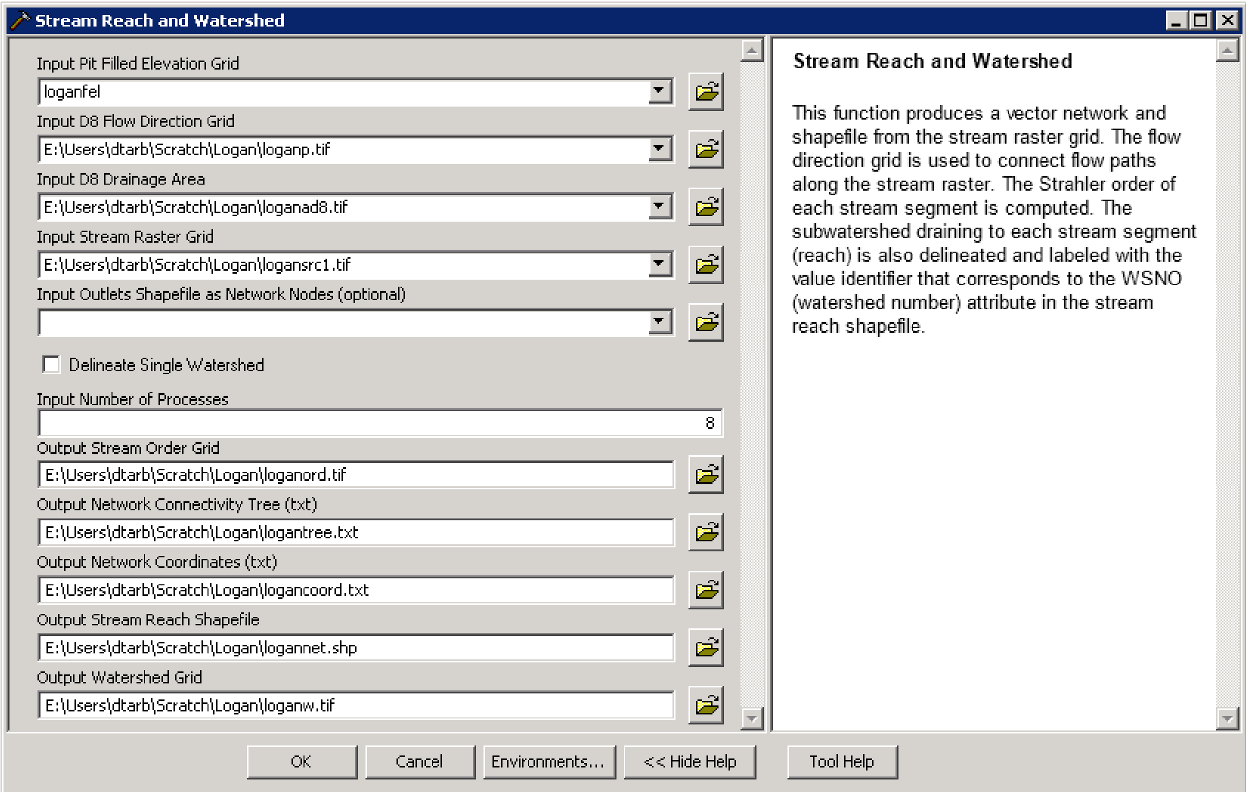
```
# Threshold
system("mpiexec -n 8 threshold -ssa loganssa.tif -src logansrc1.tif
-thresh 2000")
src1=raster("logansrc1.tif")
plot(src1)
zoom(src1)
```

The result is a grid stream network upstream of the outlet



13. This network is still only represented as a grid. To convert this into vector elements represented using a shapefile, the **Stream Reach and Watershed** function is used.

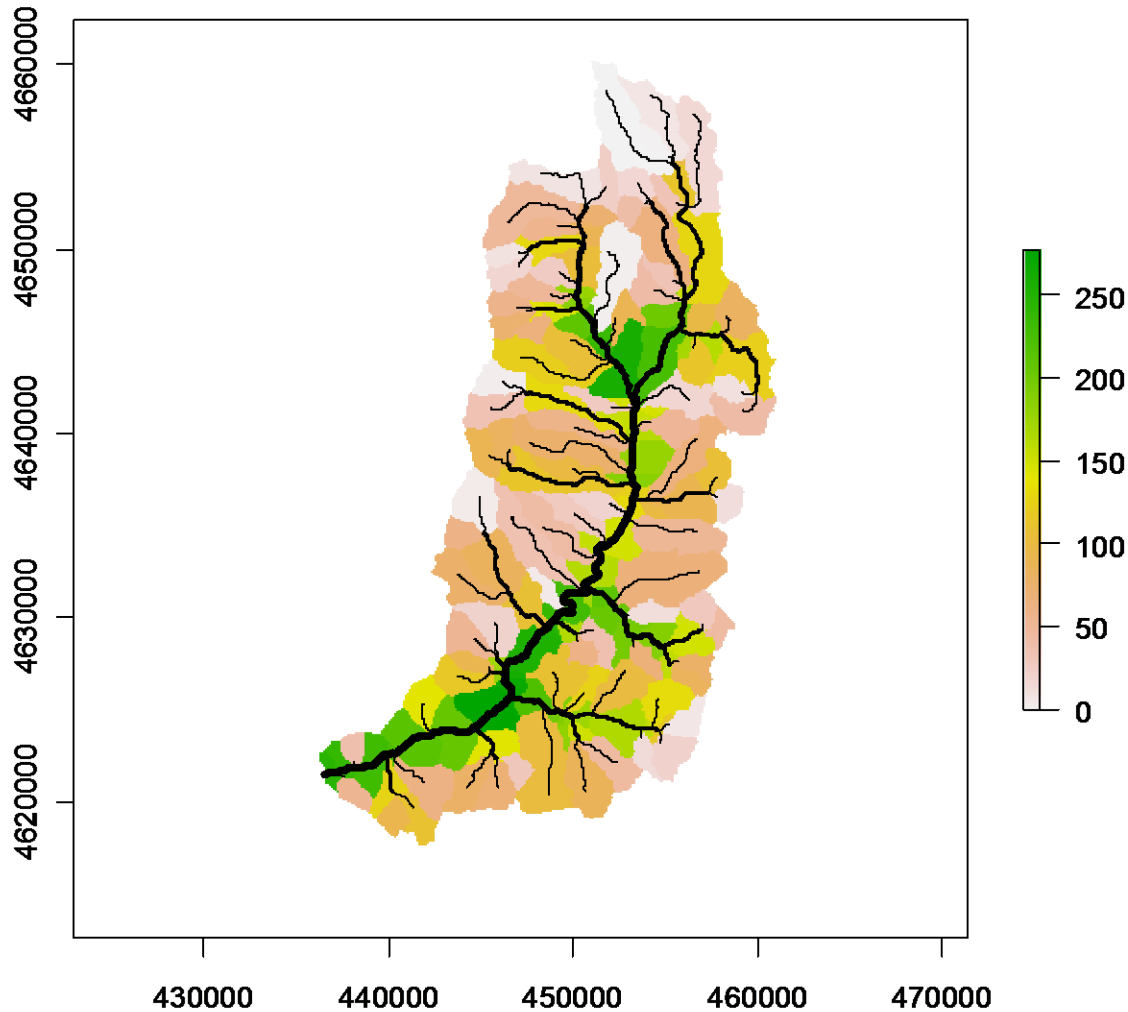
ArcMap



R

```
# Stream Reach and Watershed
system("mpiexec -n 8 Streamnet -fel loganfel.tif -p loganp.tif -ad8
loganad8.tif -src logansrc1.tif -o outlet.shp -ord loganord.tif -
tree logantree.txt -coord logancoord.txt -net logannet.shp -w
loganw.tif")
plot(raster("loganord.tif"))
zoom(raster("loganord.tif"))
plot(raster("loganw.tif"))

# Plot streams using stream order as width
snet=read.shapefile("logannet")
ns=length(snet$shp$shp)
for(i in 1:ns)
{
  lines(snet$shp$shp[[i]]$points,lwd=snet$dbf$dbf$Order[i])
}
```

The result is a number of outputs that need to be added to visualized. Here I visualized the watershed grid, loganw.tif, and stream network shapefile, logannet.shp, in R.

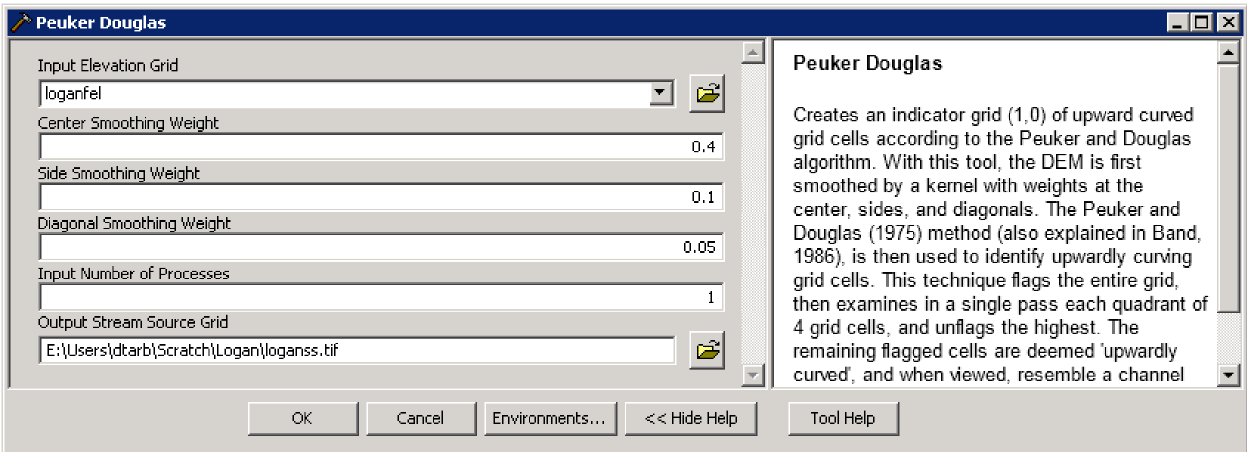


The subwatershed raster and stream network shapefile are key outputs from TauDEM. Each link in the stream network has a unique identifier that is linked to downstream and upstream links. Each subwatershed also has a unique identifier that is referenced in terms of the stream network that it drains to. This information enables construction of a subwatershed based distributed hydrologic model with flow from subwatersheds being connected to, accumulated in, and routed along the appropriate stream reaches. It is interesting to zoom in on the outlet and examine the properties of the stream network and subwatersheds near the outlet to identify how the linkages between stream links are represented and what other attributes there are for each stream link. This is easiest to do in an interactive visualization system such as ArcMap. It is a little cumbersome in R.

An important question in stream network delineation is what stream delineation threshold to use. The above used an arbitrary thresholds of 2000 grid cells. TauDEM also provides ways to do select stream delineation threshold objectively using a stream drop test following theory described in Tarboton et al. (Tarboton et al., 1991; 1992) and Tarboton and Ames (2001). This may be combined with more geomorphologically based methods for channel definition that attempt to capture topographic texture and spatial variability in drainage density. These will be illustrated next.

14. The **Peuker Douglas** function produces a "valley" network skeleton following the procedure described in their paper (Peuker and Douglas, 1975).

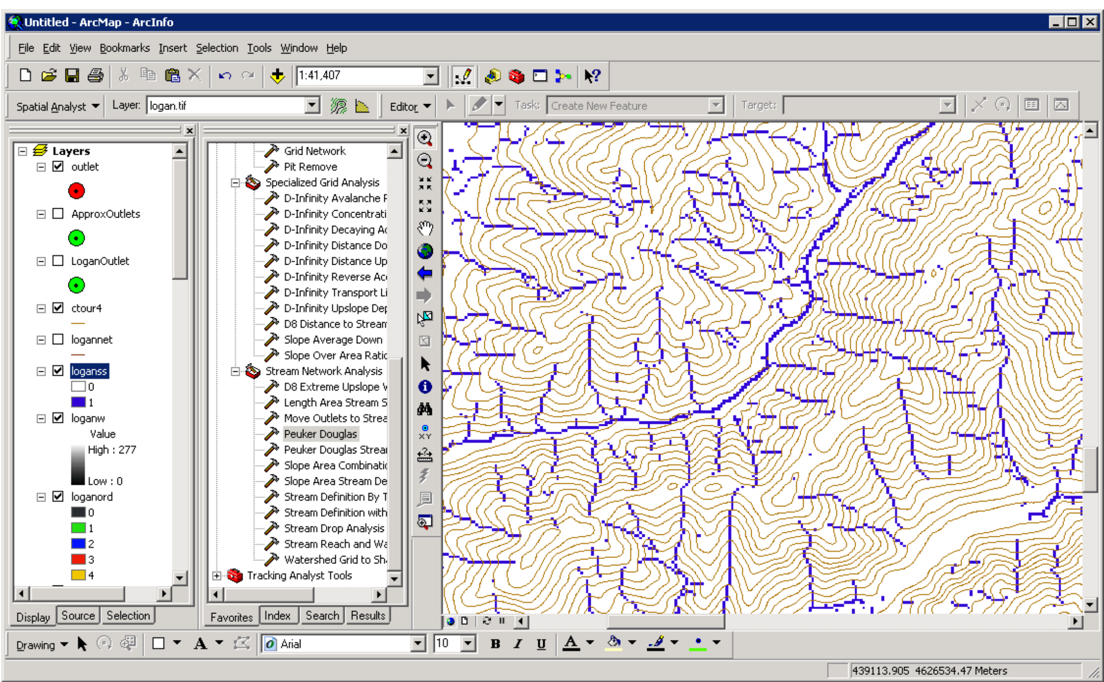
ArcMap



R

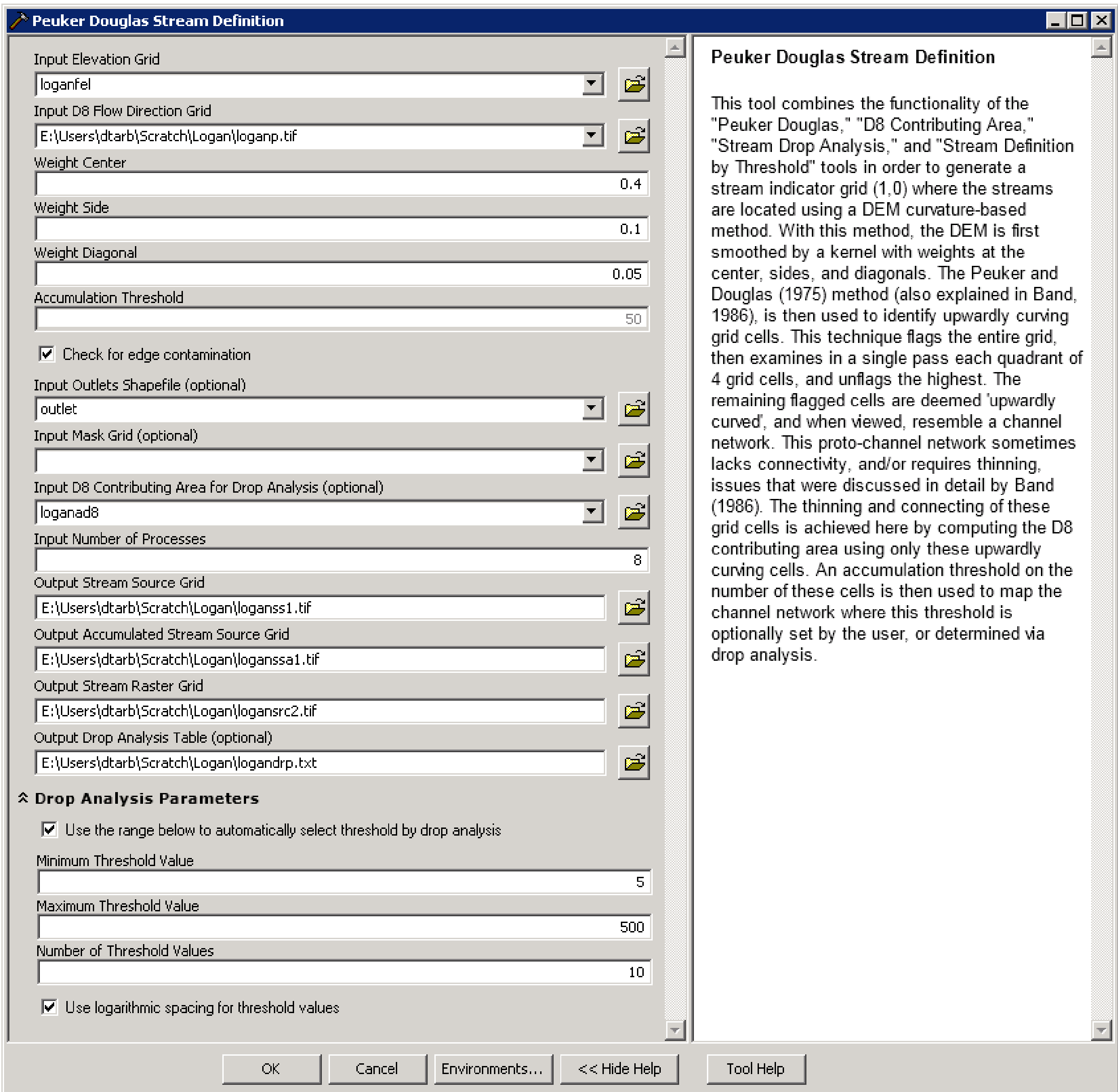
```
# Peuker Douglas stream definition
system("mpiexec -n 8 PeukerDouglas -fel loganfel.tif -ss
loganss.tif")
ss=raster("loganss.tif")
plot(ss)
zoom(ss)
```

The result, derived entirely from a local filter applied to the topography is a skeleton of a stream network illustrated below

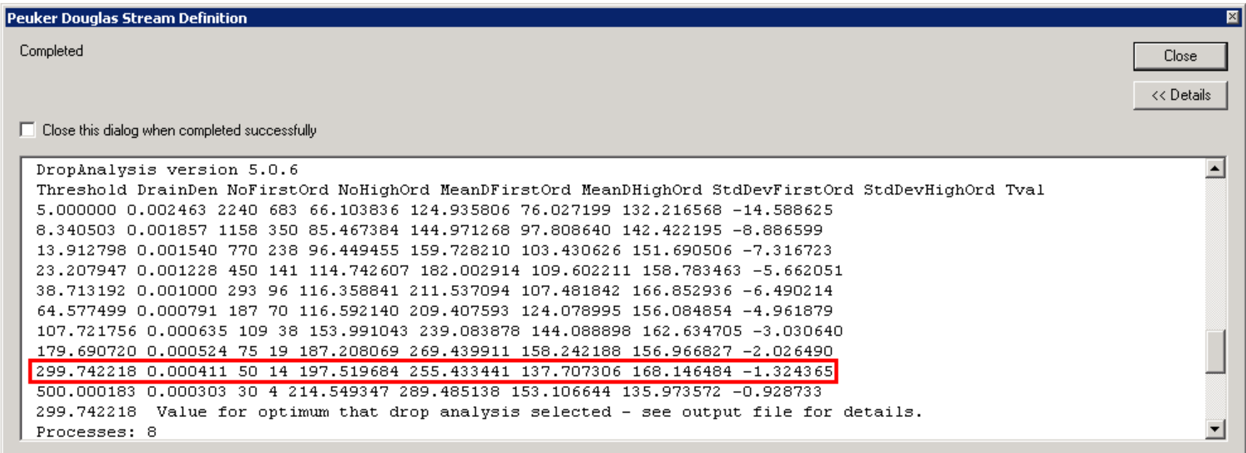


15. The **Peuker Douglas Stream Definition** function has a rather extensive set of inputs. It configures TauDEM to run Peuker Douglas, then use the Peuker Douglas stream skeleton as a weighted input to a D8 contributing area calculation. The result is then thresholded using a range of thresholds (the drop analysis thresholds at the bottom of the input) to identify the smallest threshold for which the mean stream drop of first order streams is not significantly different from the mean stream drop of higher order streams. This is the constant drop law (Broscoe, 1959), and TauDEM uses it here to identify the highest resolution stream network that complies with this law as an objective way of identifying the stream delineation threshold.

ArcMap



The output results include a table that reports the stream drop statistics for each threshold examined. This is included in the completion dialog as well as written to the drop analysis table file shown below. The last column of this gives T statistics for the differences of first and higher order streams. Using a threshold of |2| as indicating significance in this T test the threshold of 299 is chosen in this case as the objective stream delineation threshold.

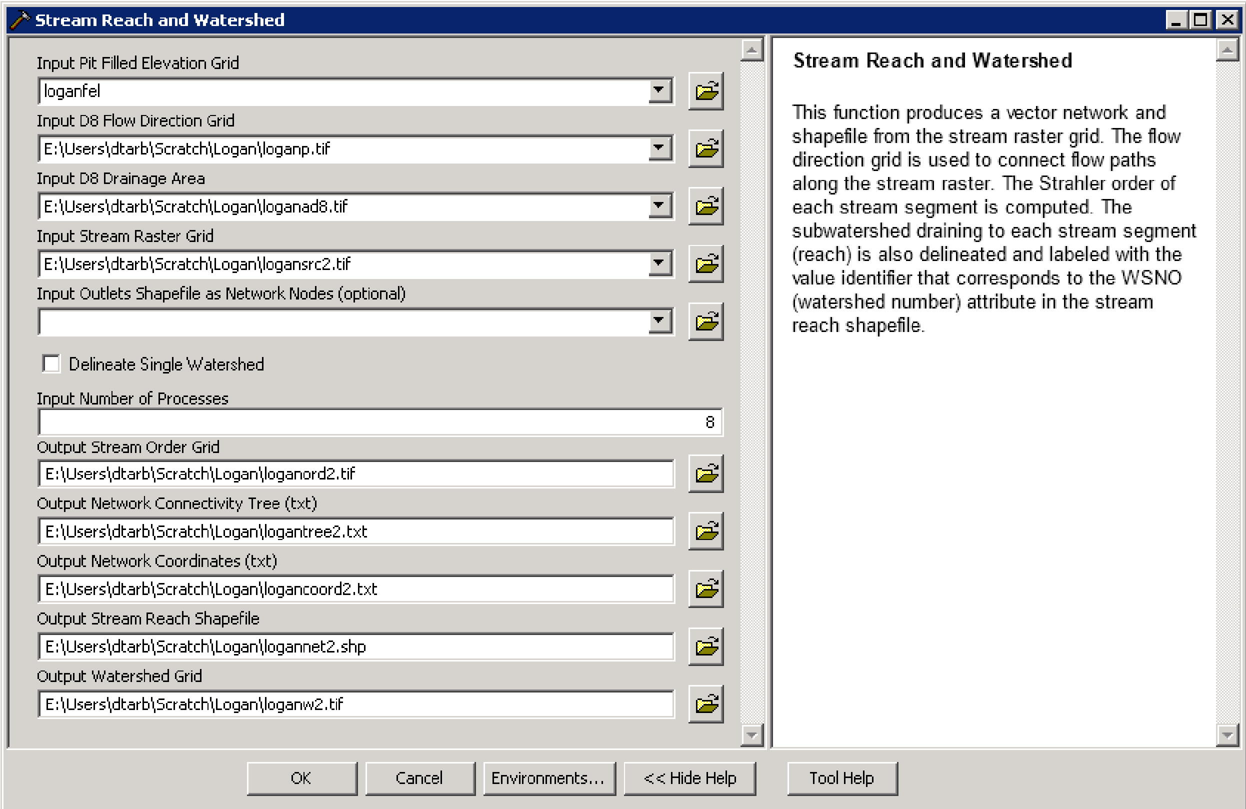


R. The ArcMap tool above ran 4 underlying TauDEM commands. The PeukerDouglas command was run earlier. Here are the next three.

```
#  Accumulating candidate stream source cells
system("mpiexec -n 8 Aread8 -p loganp.tif -o outlet.shp -ad8 
loganssa.tif -wg loganss.tif")
ssa=raster("loganssa.tif")
plot(ssa)

#  Drop Analysis
system("mpiexec -n 8 Dropanalysis -p loganp.tif -fel loganfel.tif -
ad8 loganad8.tif -ssa loganssa.tif -drp logandrp.txt -o outlet.shp -
par 5 500 10 0")

# Deduce that the optimal threshold is 300 
# Stream raster by threshold
system("mpiexec -n 8 Threshold -ssa loganssa.tif -src logansrc2.tif 
-thresh 300")
plot(raster("logansrc2.tif"))
```

16. Next the **Stream Reach and Watershed** function is used to produce a vector stream shapefile from the resulting stream raster.

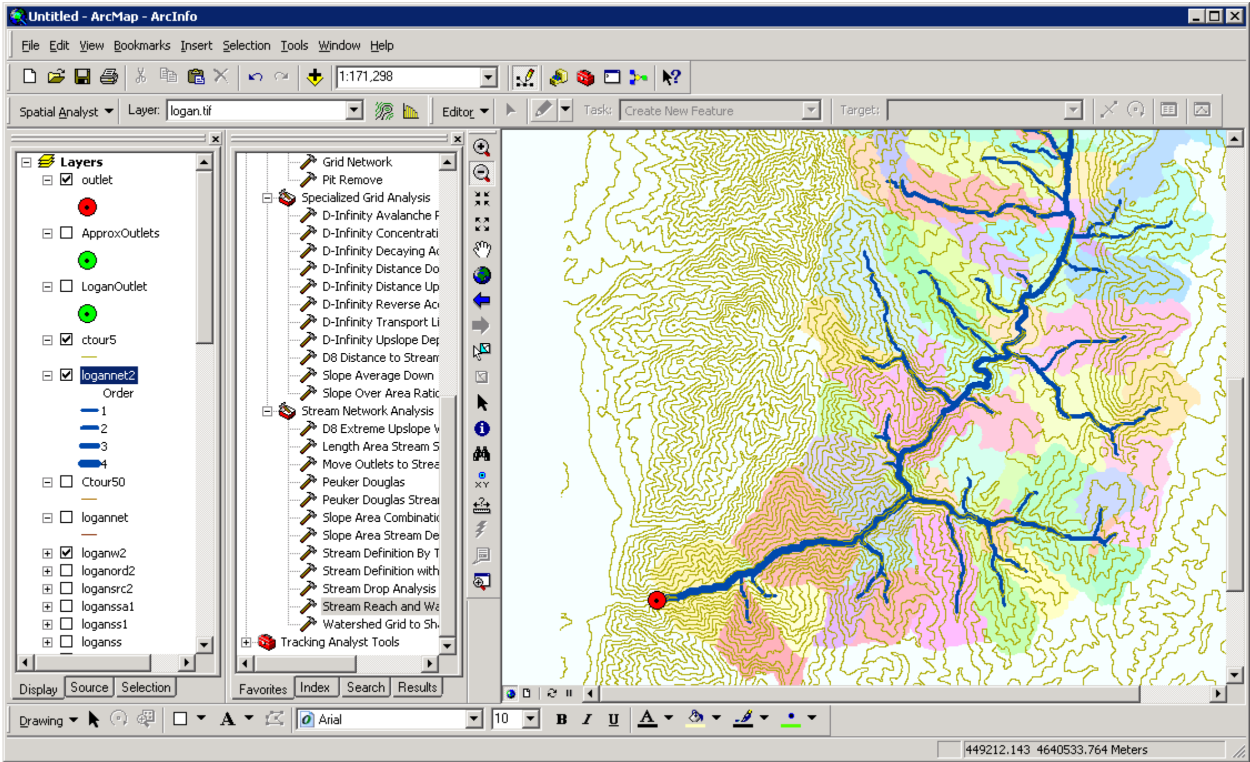
ArcMap



R

```
# Stream network
system("mpiexec -n 8 Streamnet -fel loganfel.tif -p loganp.tif -ad8
loganad8.tif -src logansrc2.tif -ord loganord2.tif -tree
logantree2.dat -coord logancoord2.dat -net logannet2.shp -w
loganw2.tif -o Outlet.shp",show.output.on.console=F,invisible=F)

plot(raster("loganw2.tif"))
snet=read.shapefile("logannet2")
ns=length(snet$shp$shp)
for(i in 1:ns)
{
  lines(snet$shp$shp[[i]]$points,lwd=snet$dbf$dbf$Order[i])
}
```

Following is an illustration of the result. Notice how the stream network has been delineated more or less consistently with the contour crenulations depicting the texture of the topography.

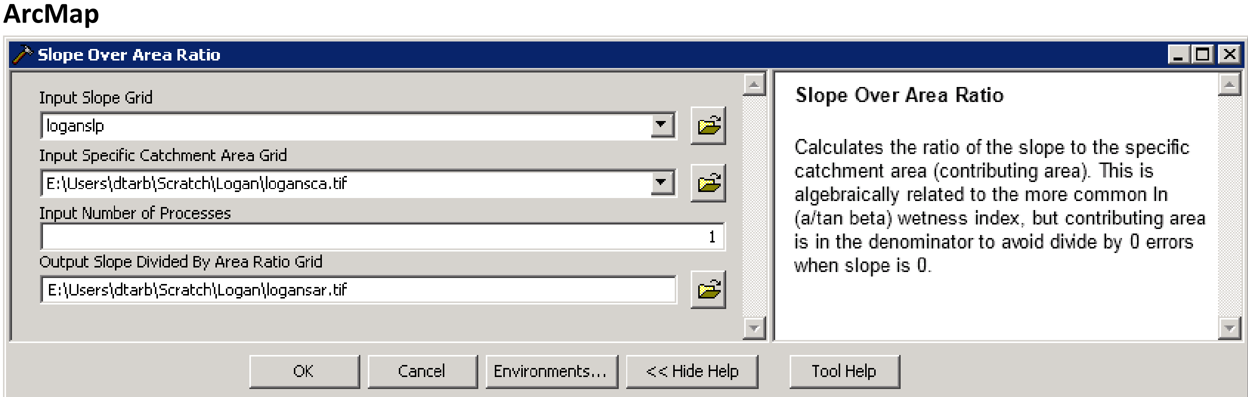


Specialized Grid Analysis using TauDEM functions

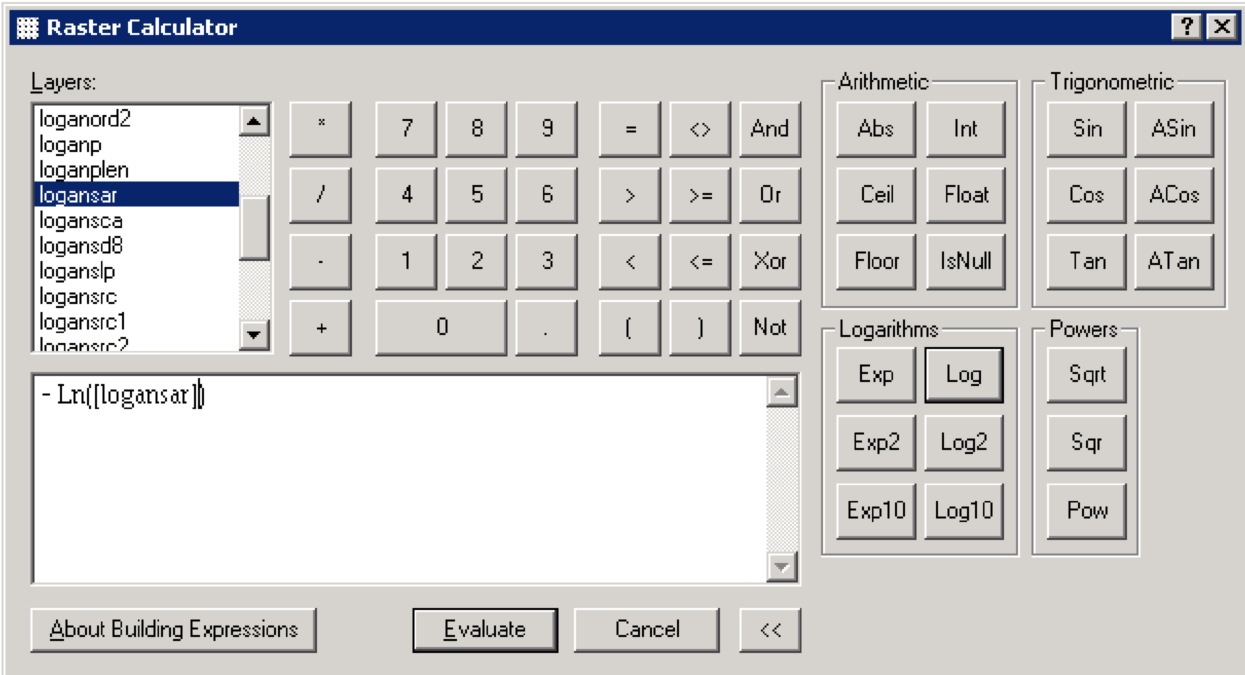
TauDEM also includes a number of specialized grid analysis functions. A few are illustrated here.

17. The TOPMODEL wetness index is defined as ls(a/S) where a is specific catchment area and S is slope (tan of slope angle). In the TauDEM outputs above a is represented by sca, the D-Infinity specific catchment area grid and S by slp, the D-Infinity slope. sca is alreay in length units (the same units as grid cell size). TauDEM has a function to evaluate sar=S/a. This is done to leave to the user the choice as to how to handle grid cells that have S=0. Wetness index is then -ln(sar).

ArcMap



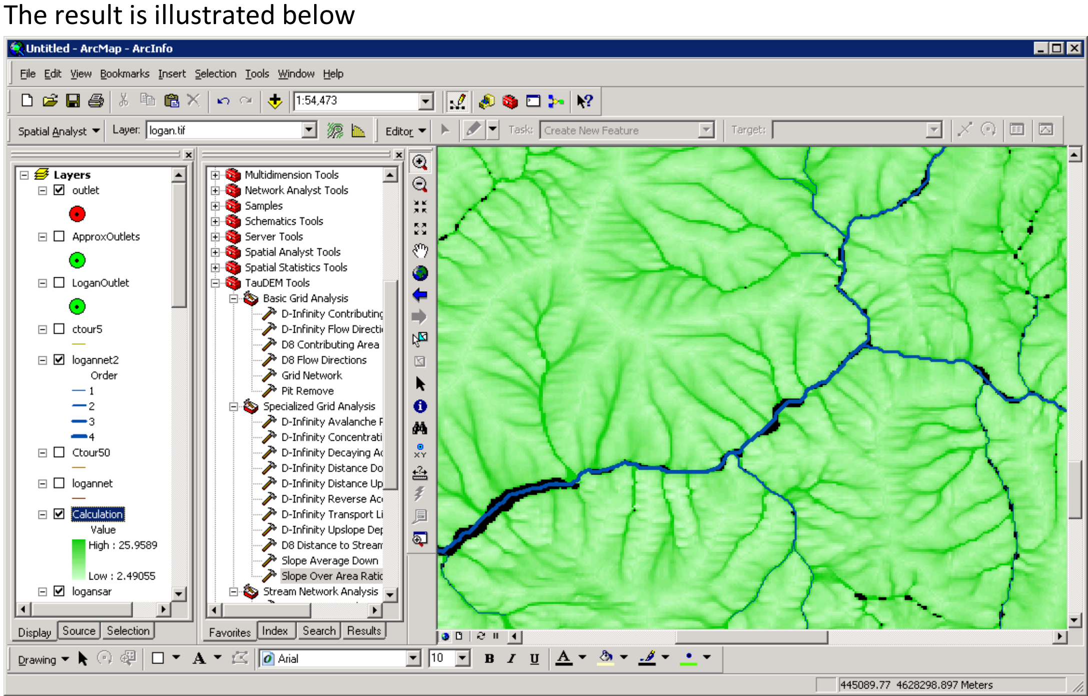
Wetness index is evaluated using the ArcMap Raster Calculator



R

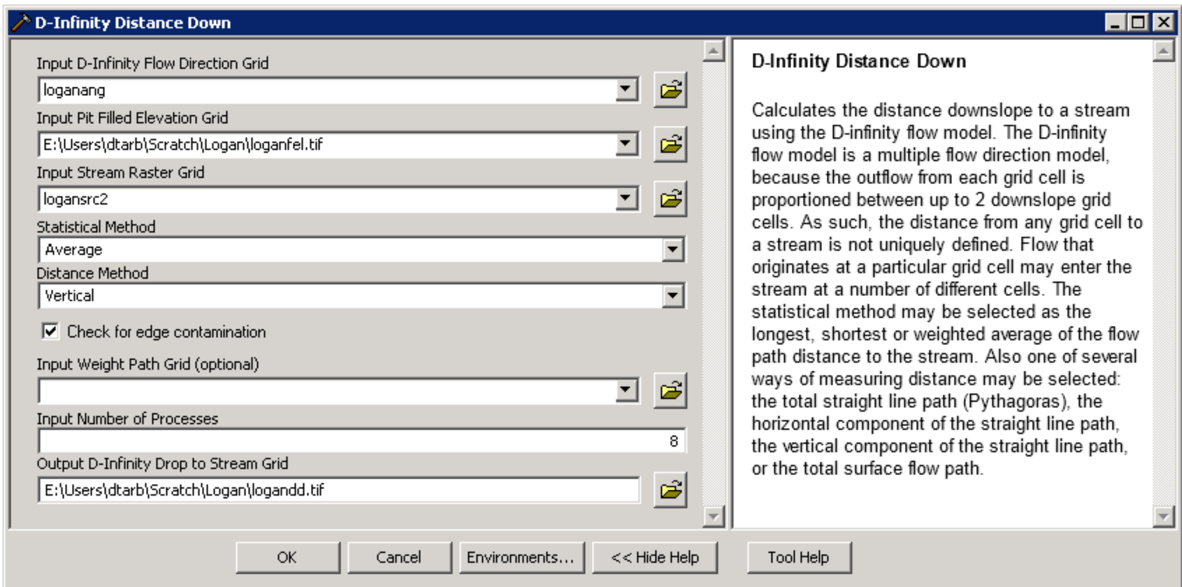
```
# Wetness Index
system("mpiexec -n 8 SlopeAreaRatio -slp loganslp.tif -sca
logansca.tif -sar logansar.tif", show.output.on.console=F,
invisible=F)
sar=raster("logansar.tif")
wi=sar
wi[,]=-log(sar[,])
plot(wi)
```

The result is illustrated below



18. The **D-Infinity Distance Down** function computes the distance to streams (or any designated target grid) a number of different ways

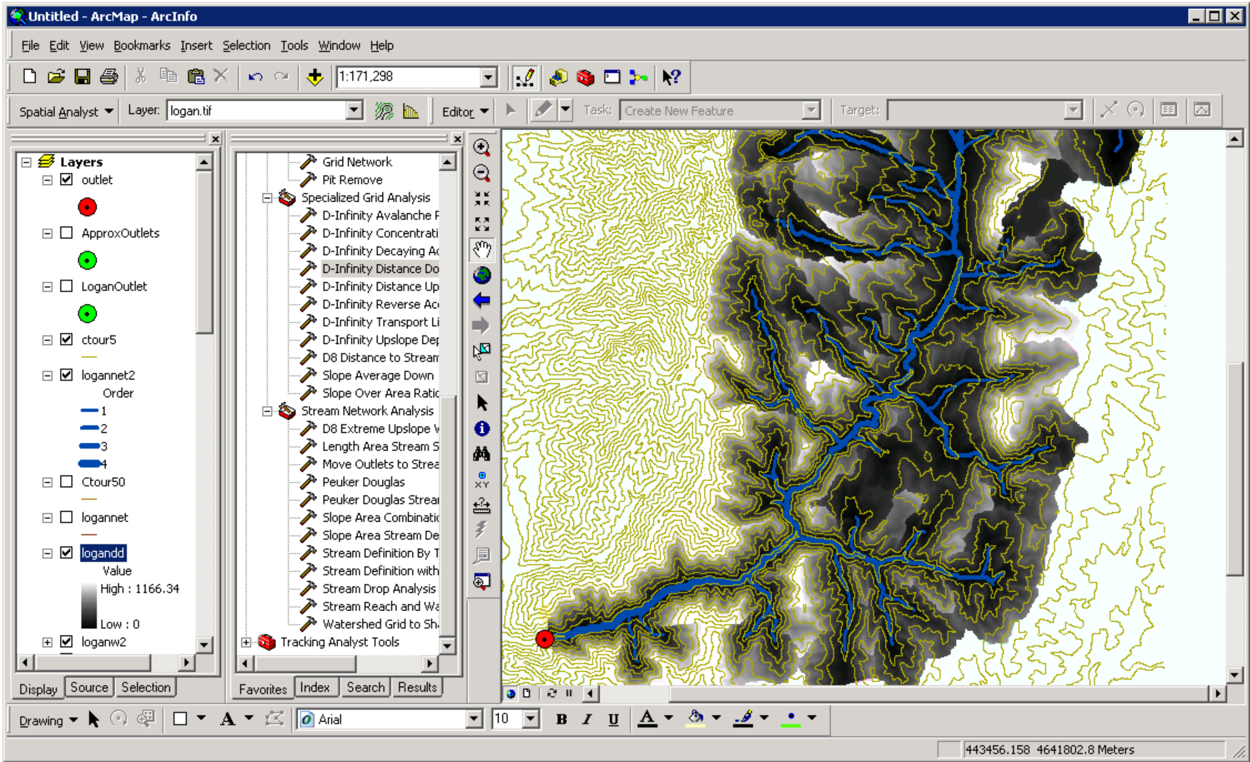
ArcMap



R

```
# Distance Down
system("mpiexec -n 8 DinfDistDown -ang loganang.tif -fel
loganfel.tif -src logansrc2.tif -m ave v -dd
logandd.tif",show.output.on.console=F,invisible=F)
plot(raster("logandd.tif"))
```

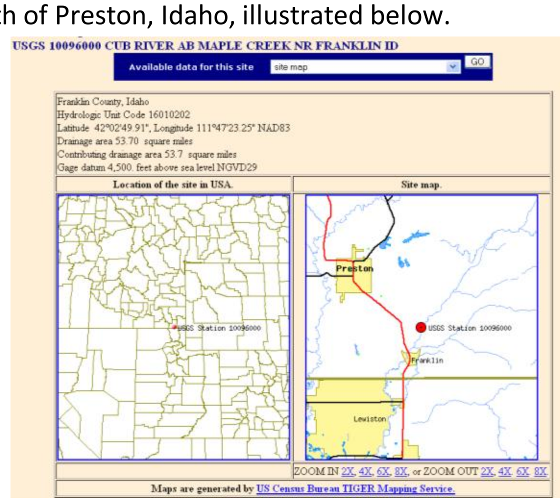
By selecting **-m ave v** as the distance method the result is the average vertical drop from each point, to a point on the stream as illustrated below.



There are many other options for distance methods that are described in the help file http://hydrology.usu.edu/taudem/taudem5.0/TauDEM_Tools.chm and command line guide http://hydrology.usu.edu/taudem/taudem5.0/TauDEM5CommandLineGuide.pdf that you could experiment with if you want to.

Appendix 1. Downloading DEM data from the USGS Seamless data server

This appendix illustrates the process of downloading and projecting DEM data from the USGS Seamless data server, for the Cub River watershed as it drains to the location of a USGS streamflow station #10096000 located just north of Preston, Idaho, illustrated below.

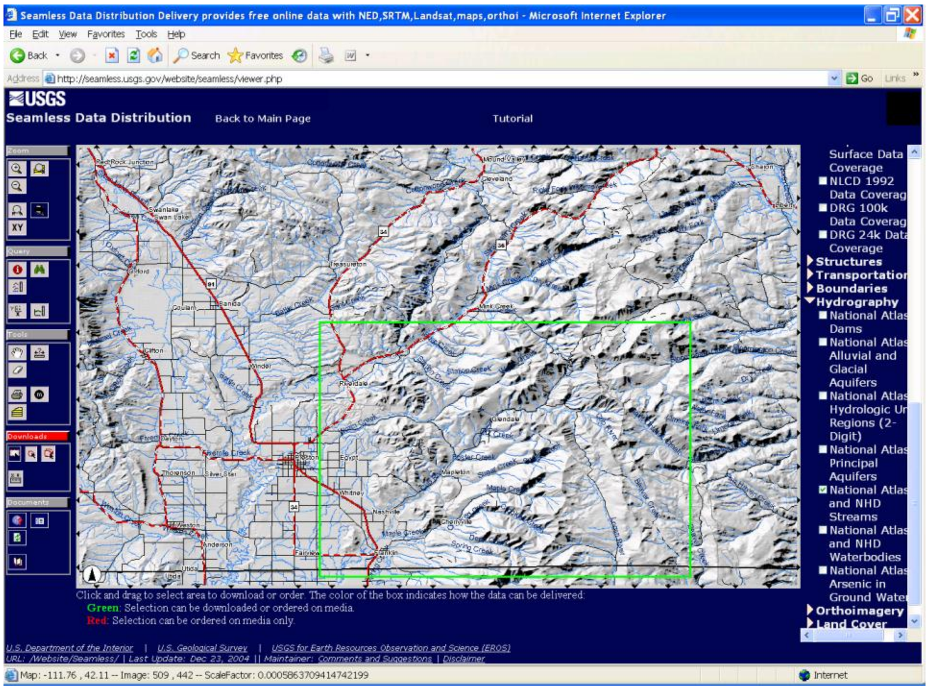


The USGS Seamless data server was used to obtain a National Elevation Dataset DEM.

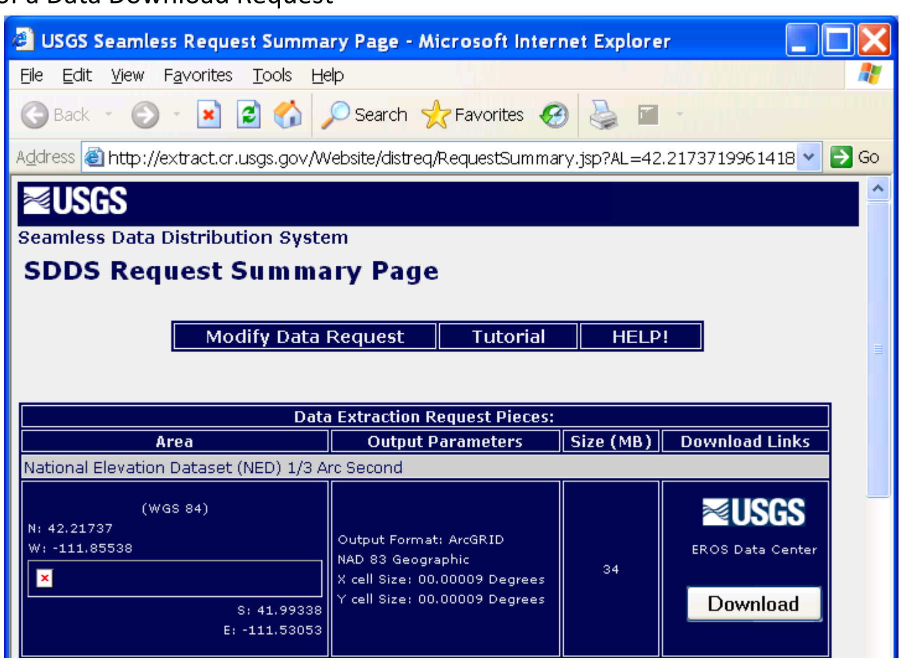
The steps followed were

1. Access USGS Seamless data server http://seamless.usgs.gov/
2. Click view and download United States Data
3. Zoom to the area of interest. Activate layers on the right to help identify area of interest.
4. Define a download region that covers the area of interest.
5. Modify the data request to comprise the data sets (parameters) that you want to obtain
6. Download the data. I selected the 1/3 arc second National elevation dataset DEM (≈ 10 m grid)

Screen Image of the area that I selected

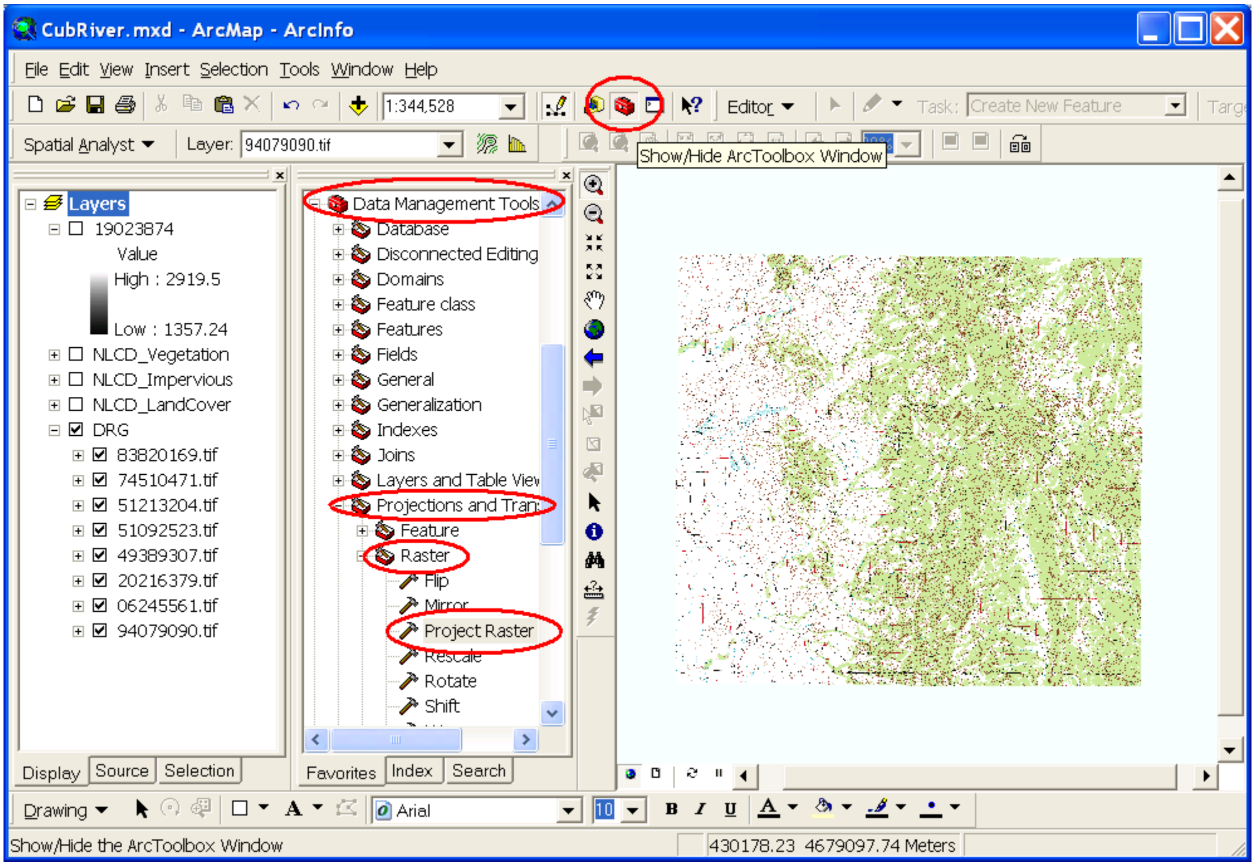


Screen Image of a Data Download Request

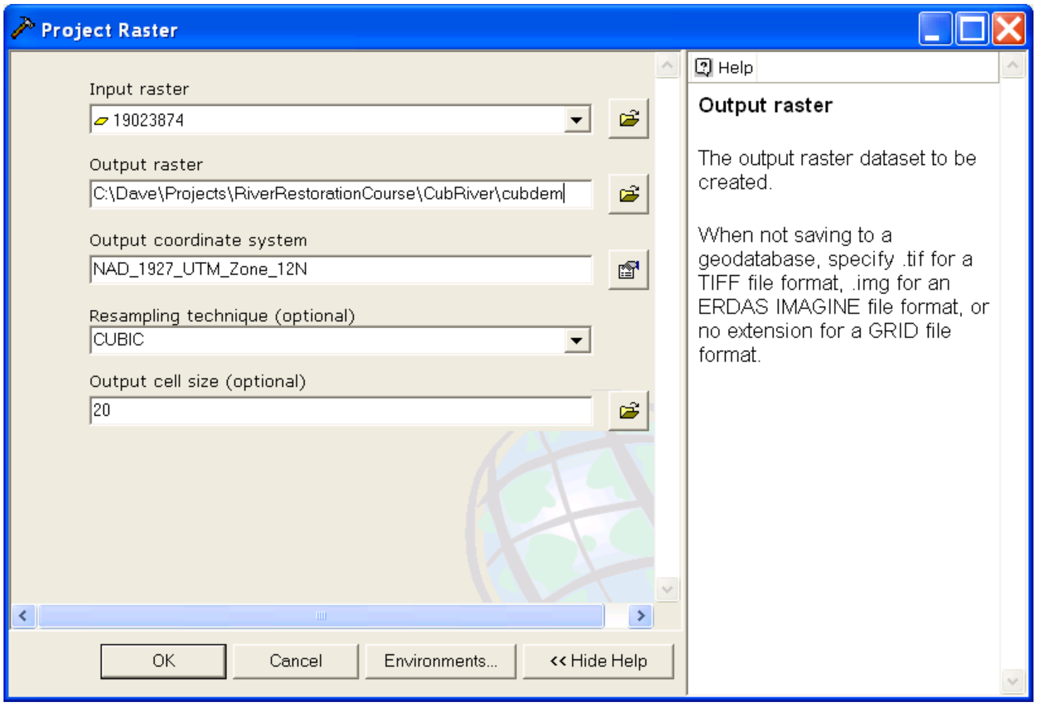


Projecting the Digital Elevation Model data

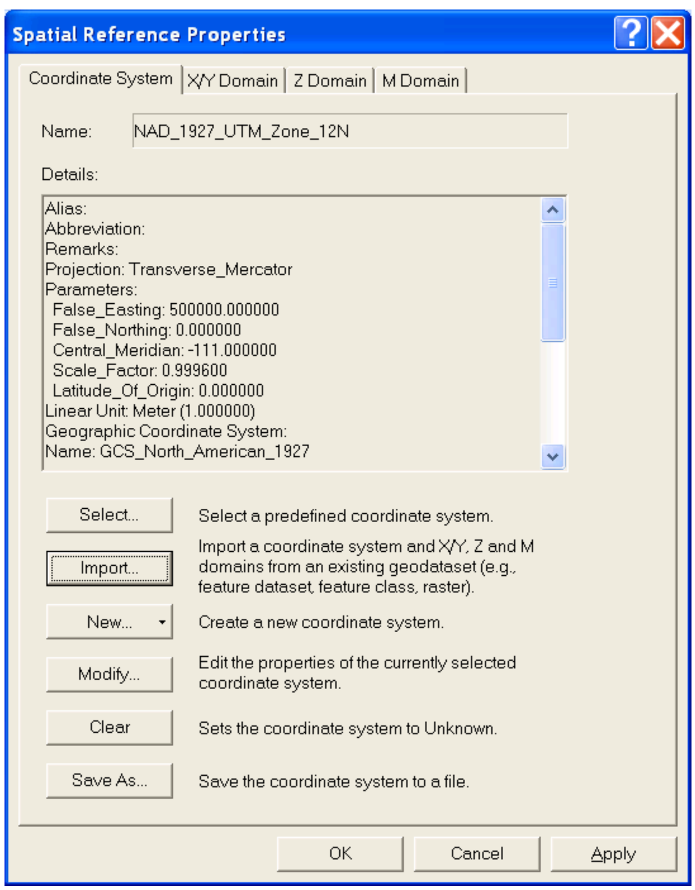
The Digital Elevation Model grid from the Seamless Data Server is in Geographic Coordinates. Projected data should be used when working with TauDEM because TauDEM uses grid dimensions (cell size) in its length and slope calculations and these will be incorrect if they are not consistent in E-W and N-S directions and in the same units as the vertical units of the DEM. The DEM from the USGS was added to ArcMap. Then the ArcToolBox Project Raster tool was used to project this data. The ProjectRaster Tool is found within Data Management Tools / Projections and Transformations / Raster.



In the Project Raster dialog that opens specify the input raster as the National Elevation Dataset DEM that was unzipped from the download. Name the output raster something convenient. Here I used "cubdem". Click on the button next to Output coordinate system to open the Spatial Reference Properties dialog.



At this Spatial Reference Properties dialog click "Select" and navigate to the NAD_1927_UTM_Zone_12N projection being used as the standard spatial reference system for this exercise. Click OK.



Back at the Project Raster dialog set the resampling technique to CUBIC (I have found by experience that this works best for DEMs) and set the output cell size to 20 m. The raw data in this case is at 1/3 arc second which is roughly 10 m. 20 m cell size is undersampling this a bit. Click OK. A processing dialog box should appear and after a few seconds indicate completion of the projection of the DEM. The DEM data has now been projected. The result is named 'cubdem'

References

Broscoe, A. J., (1959), "Quantitative analysis of longitudinal stream profiles of small watersheds," Office of Naval Research, Project NR 389-042, Technical Report No. 18, Department of Geology, Columbia University, New York.

Peuker, T. K. and D. H. Douglas, (1975), "Detection of surface-specific points by local parallel processing of discrete terrain elevation data," Comput. Graphics Image Process., 4: 375-387.

Tarboton, D. G. and D. P. Ames, (2001), "Advances in the mapping of flow networks from digital elevation data," World Water and Environmental Resources Congress, Orlando, Florida, May 20-24, ASCE, http://www.engineering.usu.edu/dtarb/asce2001.pdf.

Tarboton, D. G., R. L. Bras and I. Rodriguez-Iturbe, (1991), "On the Extraction of Channel Networks from Digital Elevation Data," Hydrologic Processes, 5(1): 81-100.

Tarboton, D. G., R. L. Bras and I. Rodriguez-Iturbe, (1992), "A Physical Basis for Drainage Density," Geomorphology, 5(1/2): 59-76.