

SHORT COURSE FOR IDENTIFICATION OF SINKHOLES FROM LIDAR IMAGERY

15th Sinkhole Conference and 3rd Appalachian Karst Symposium
National Conservation Training Center,
Shepherdstown, West Virginia
April 2nd 2018

By: Daniel H. Doctor, U.S. Geological Survey
and

John Wall, Ph.D. candidate, North Carolina State University



INTRODUCTION

This short course will introduce users to several tools and techniques for identifying sinkholes and other closed depressions from lidar-derived digital elevation models (DEMs) within ESRI ArcGIS software. The short course is designed to take the user through a series of steps that are typical for visualization and analysis of DEMs for finding and morphometrically characterizing closed depressions.

Topics to be covered will include:

- 1) Semi-automated methods for extracting closed depression features from DEMs using filling and contour-tree algorithms
- 2) Hydrologic conditioning of lidar-derived DEMs for removal of false closed depressions along 'digital dams', such as roads and railways
- 3) Implementation of flow-routing algorithms for generation of vector stream networks from DEMs
- 4) Optimized visualization of lidar-derived elevation data for the identification of karst landscape features

Additional skills to be learned will include:

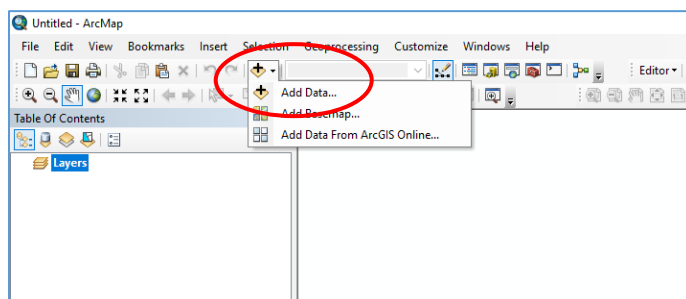
- 1) Enhanced visualization of lidar elevation datasets
- 2) Access to publicly-available lidar data
- 3) Using GIS software to extract closed depression polygons and stream vector data from lidar elevation models
- 4) Export and visualization of lidar data to Google Earth

SETTING UP A PROJECT IN ArcMAP®

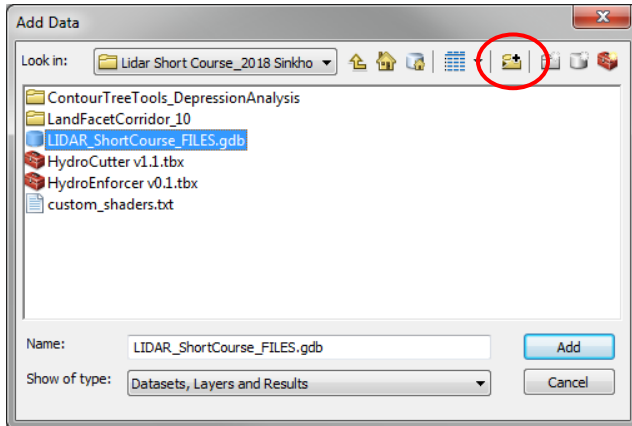
On the desktop there is folder named *LidarShortCourse_2018SinkholeConference* that contains a number of data files. Please open this folder.

NOTE: It is good practice to refrain from including spaces within path names. ArcGIS was historically built with UNIX practices in mind; thus, it is also good practice to 1) operate as close to the C:\ of your machine as possible; 2) refrain from using spaces between names and within your file path; 3) do not using illegal characters (e.g. !, @, %); and 4) keep file names to less than 13-characters in length. While there are work-arounds, such as placing files in a file geodatabase to get around the 13-character rule, these have been reported to cause issues with the tools used in this short course.

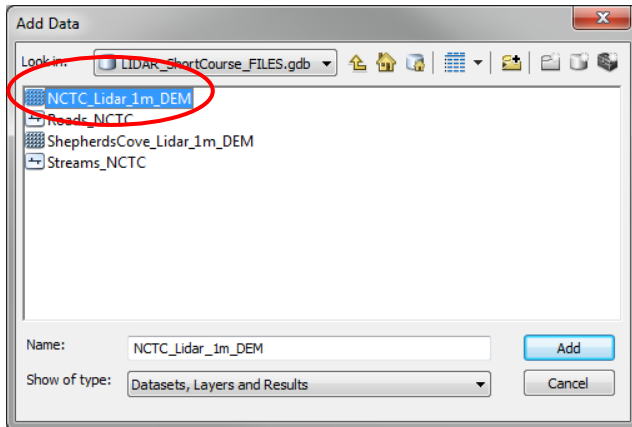
To begin, launch ArcMap, and select the plus icon to add data; select **Add Data...**:



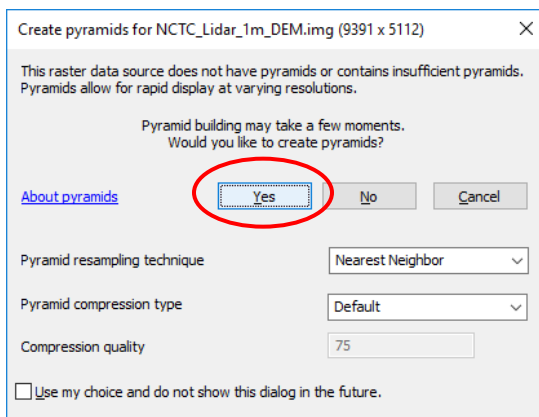
Select the **Connect to Folder** icon (folder with plus sign), and then navigate to the folder where the data are stored on your hard disk. Once this is set, you will not need to return to it to add new data:



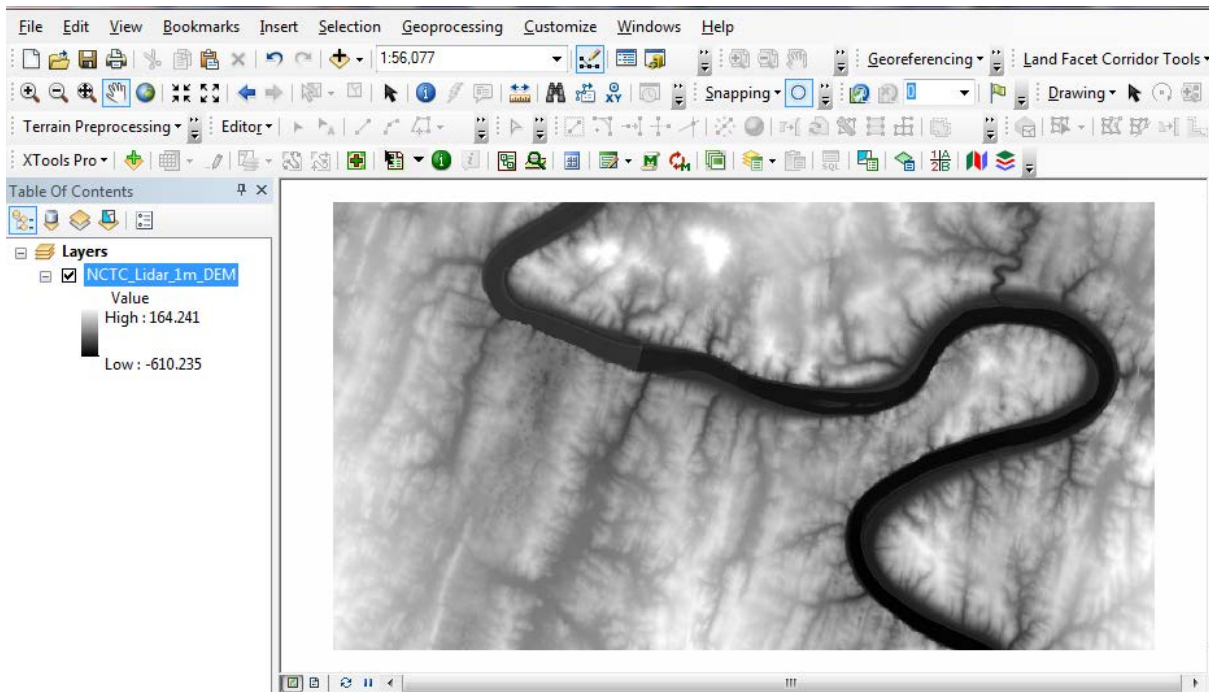
Once connected to the folder *LidarShortCourse_2018SinkholeConference* containing the data files, open the *LIDAR_Short Course_FILES.gdb* file geodatabase, and add the file *NCTC_Lidar_1m_DEM.img*. This raster is a digital elevation model (DEM) created from lidar with 1 m horizontal resolution, and 9 cm vertical Root Mean Squared Error (RMSE) which has been gridded to a 1 m resolution image:



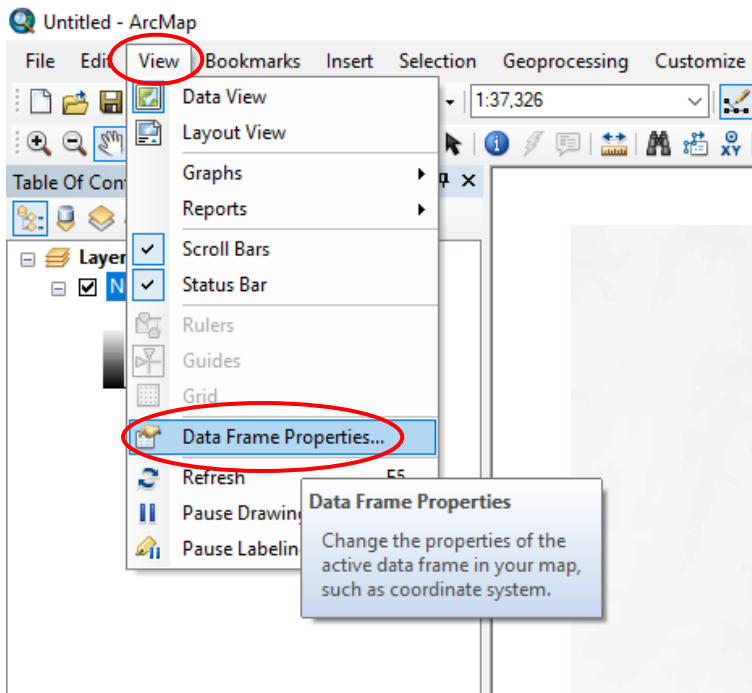
A dialog box will open asking if you'd like to create pyramids—a method of visualizing raster data quickly depending on map scale—click **Yes**:



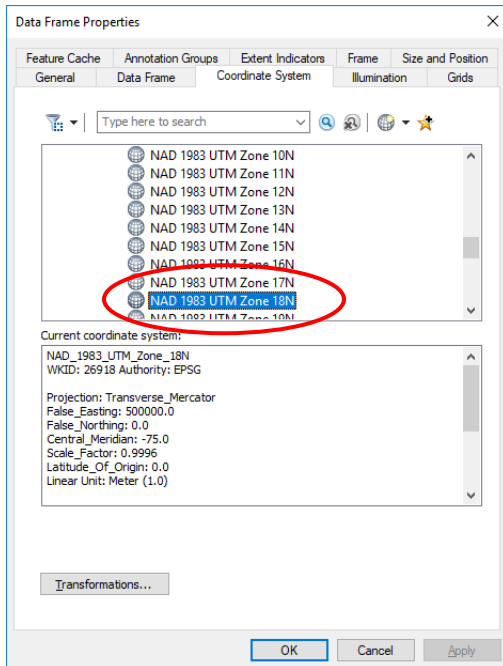
After a minute or so, the pyramid building process should be finished, and the DEM image should be placed into the **Data Frame** named *Layers* of ArcMap. You should see something similar to that below:



Check the Data Frame Properties (**View** → **Data Frame Properties**), and inspect the projection of the Data Frame, by clicking on the **Coordinate System** tab:



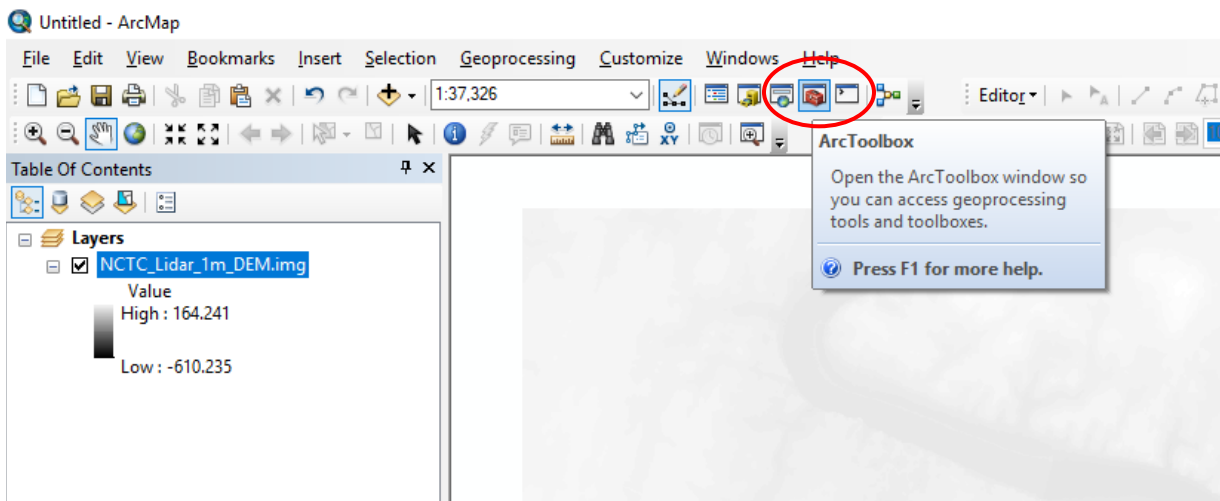
You will see that the projection of the Data Frame was set to the same projection of the first data set added to the project (i.e., the projection of the DEM) at North American Datum 1983, UTM Zone 18 North (WKID: 26918):



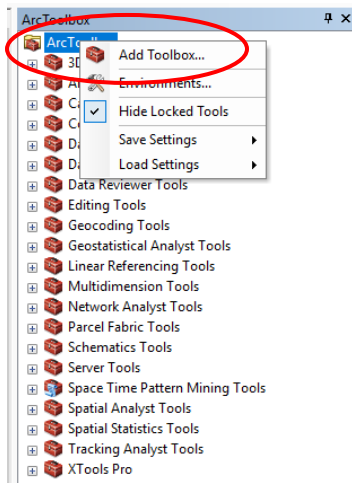
NOTE: It is crucial that every additional data set added to an ArcMap project has a datum and projection that matches the horizontal datum and projection of the Data Frame, otherwise errors may be introduced during editing.

ADD CUSTOM TOOLBOXES

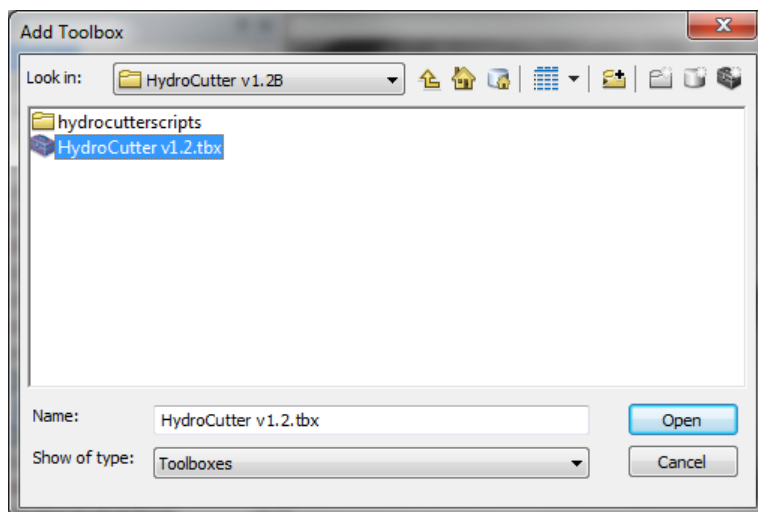
Open **ArcToolbox** to add the first custom toolbox for this short course called *HydroCutter*:



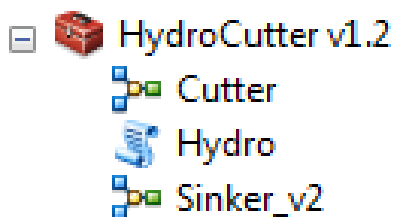
Right click on the **ArcToolbox** icon at the top of the list, and select **Add Toolbox...**:



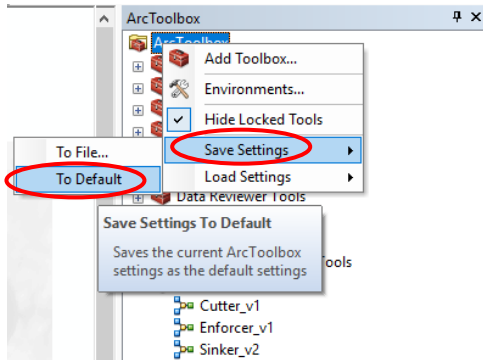
Navigate to the *LidarShortCourse_2018SinkholeConference* folder on your desktop and open the folder named *HydroCutter v1.2B*. Inside you will find the *HydroCutter v1.2.tbx*. Select the toolbox and click **Open** to add it:



Once added, expand the toolbox and you will see three tools: *Cutter*, *Hydro*, and *Sinker_v2*.



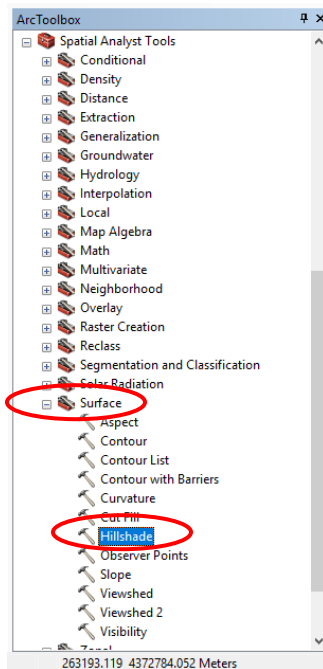
To keep the toolbox always available when you open ArcMap, right-click again on the **ArcToolbox** icon at the top of the list and select **Save Settings** → **To Default** .



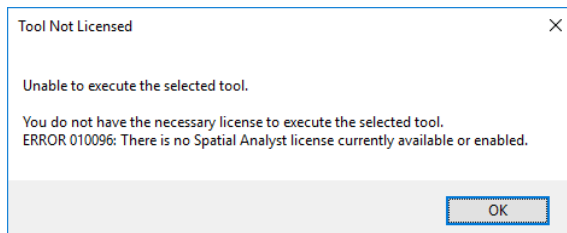
Repeat the process for the *Contour Tree Tools Public* toolbox, which will have the tools *Extract Sink* and *Identify Depression Hierarchy* in it. We will use these custom tools later on during the course.

CREATE A HILLSHADE FROM THE DEM

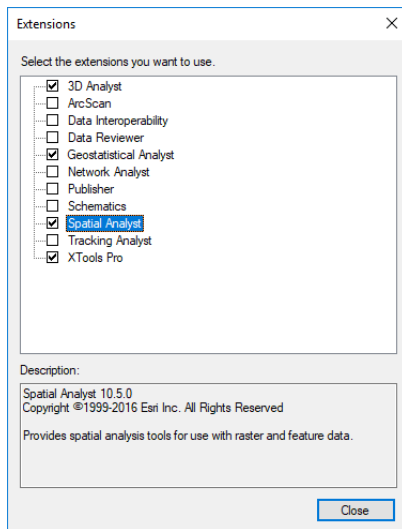
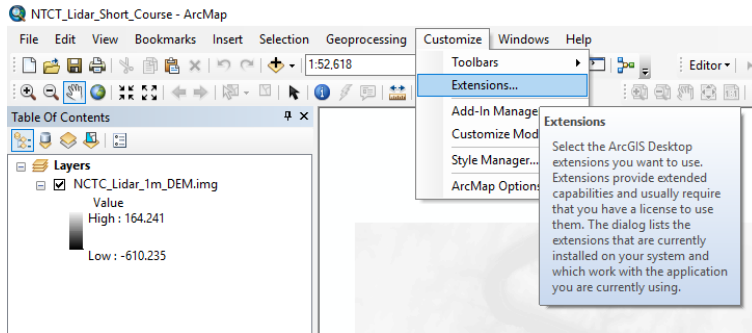
In **ArcToolbox**, navigate to **Spatial Analyst Tools** → **Surface** and click on the *Hillshade* geoprocessing tool.



If you receive the error below, you may need to enable to the **Spatial Analyst** extension first:

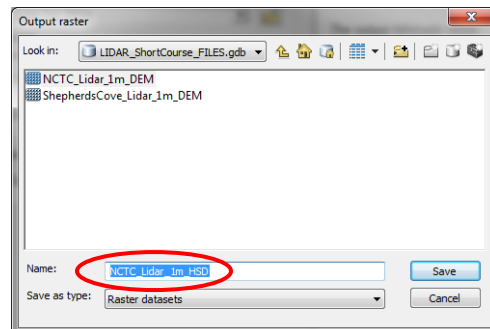
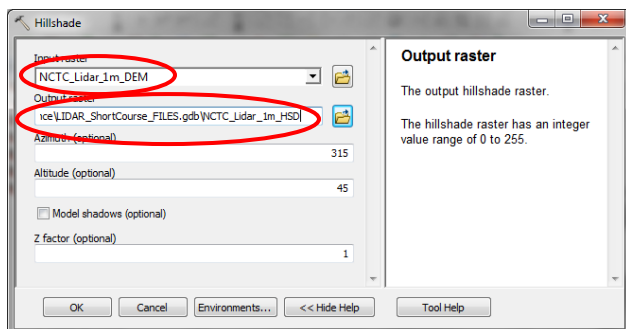


To enable the **Spatial Analyst** extension, go to **Customize** → **Extensions...** and check the box next to **Spatial Analyst** along with any other extensions you may wish to use in the future.

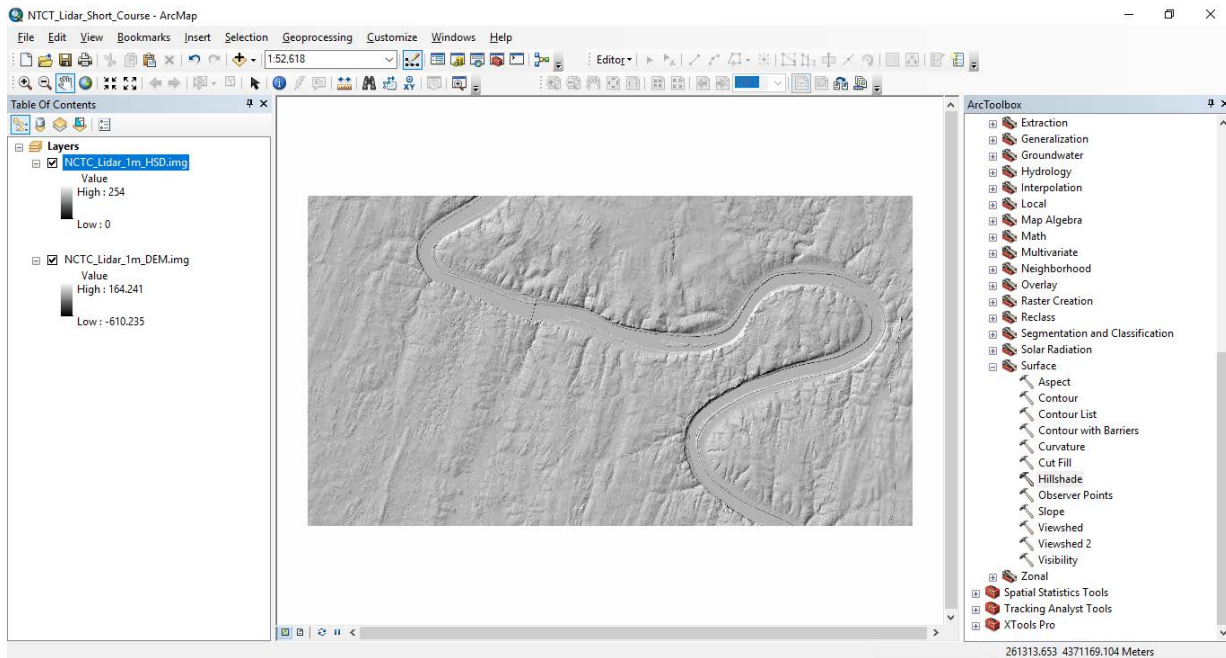


If the **Spatial Analyst** extension was already enabled, you should see the input boxes for the *Hillshade* tool. Use the drop down menu under Input Raster to select *NCTC_Lidar_1m_DEM*. Click on the Navigate to Data button (folder icon) and navigate to *Lidar Short Course_FILES.gdb* on the desktop where you will save the output raster as *NCTC_Lidar_1m_HSD*.

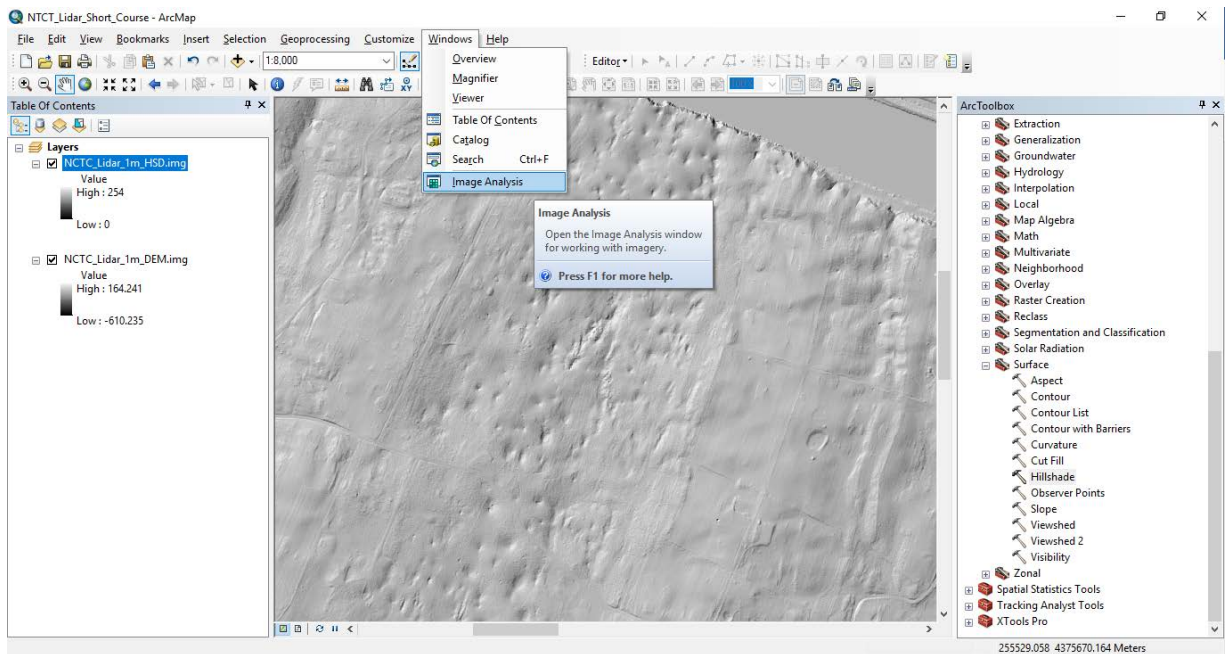
Leave the default values for **Azimuth**, **Altitude**, and **Z factor** as they are. The **Azimuth** and **Altitude** values set the synthetic illumination direction of the elevation surface, and the **Z Factor** sets the vertical exaggeration. **NOTE:** if the vertical and horizontal units of the DEM are not the same, then the **Z factor** may need to be adjusted, i.e. to convert feet to meters or vice versa. For the *NCTC_Lidar_1m_DEM*, both the horizontal and vertical units are metric.



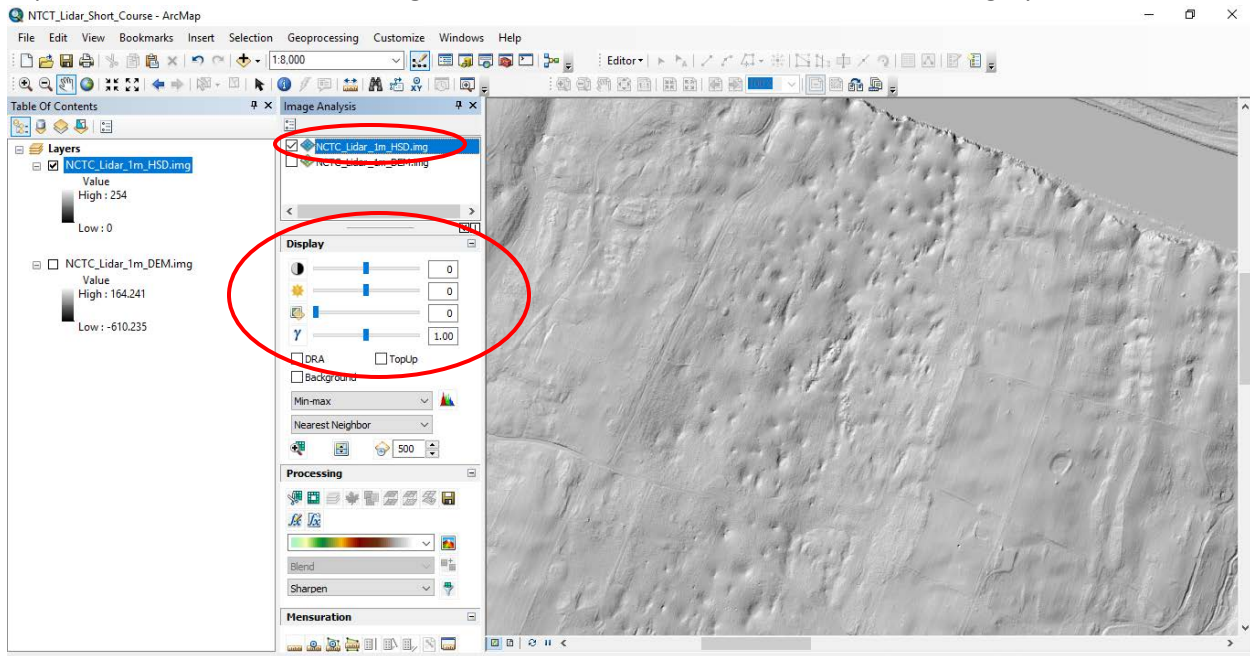
Once the *Hillshade* processing is complete, you should see a resulting image like this:



Zoom in to the area of the image with many closed depressions, and open the **Image Analysis** window:



Experiment with the **Contrast, Brightness, and Gamma** sliders to enhance the imagery:



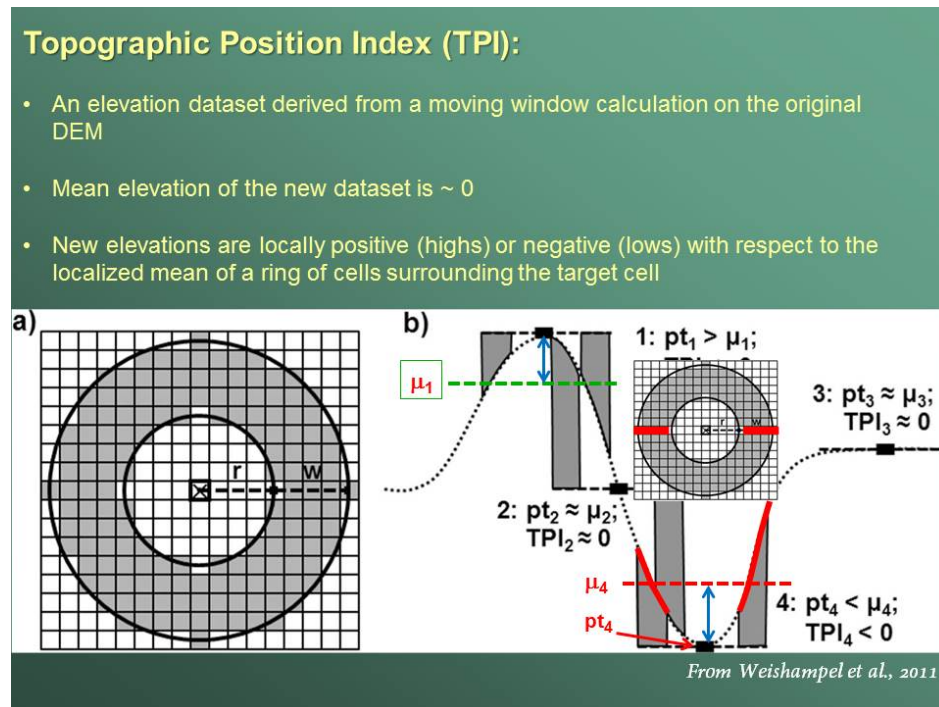
CREATE A TOPOGRAPHIC POSITION INDEX (TPI) IMAGE FROM THE DEM

The *Land Facet Corridor Analysis* tools for ArcMAP v 10.x should already be loaded into your ArcMAP tool menu. This tool is freely available at http://www.jennessent.com/arcgis/land_facets.htm, where you will find instructions for loading the tool onto your own system. The image of the website is below:

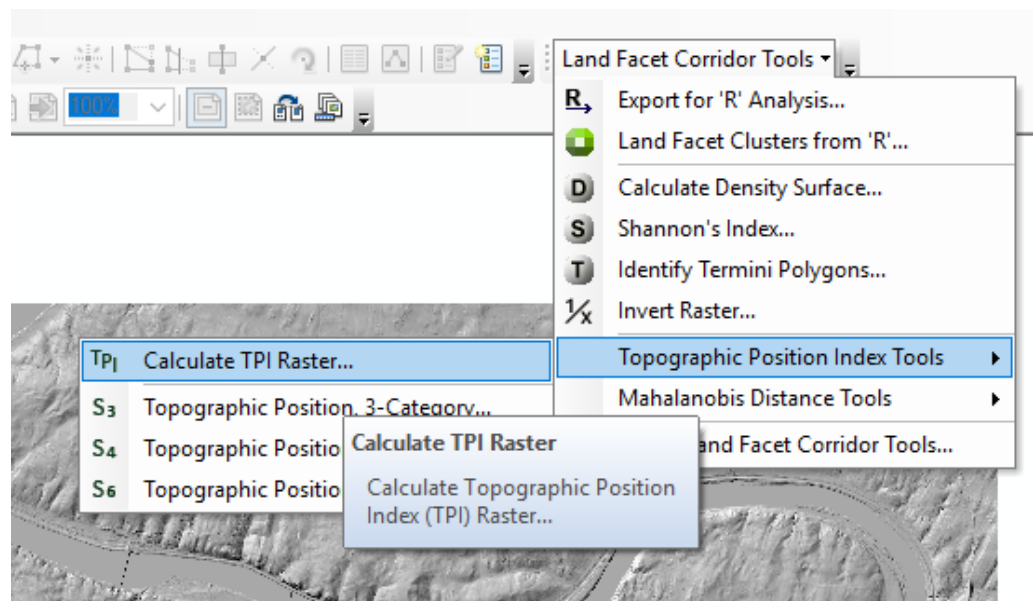


For this course, the *Topographic Position Index (TPI)* calculator is the only tool we will use from this toolset. The TPI is a measure of local relief and is calculated as a moving-window operation across the entire DEM in which the mean elevation of a region surrounding each cell is subtracted from the elevation of the central cell.

Below is a diagram illustrating the TPI calculation:

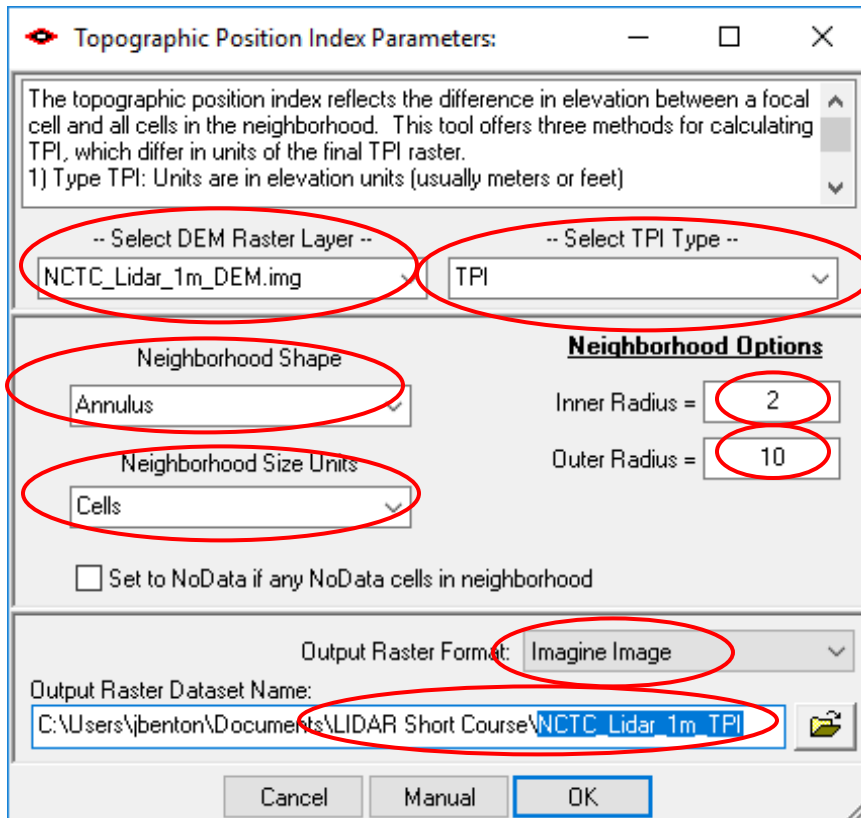


Navigate to the tool from **Land Facet Corridor Tools** → **Topographic Position Index Tools** → **Calculate TPI Raster**:

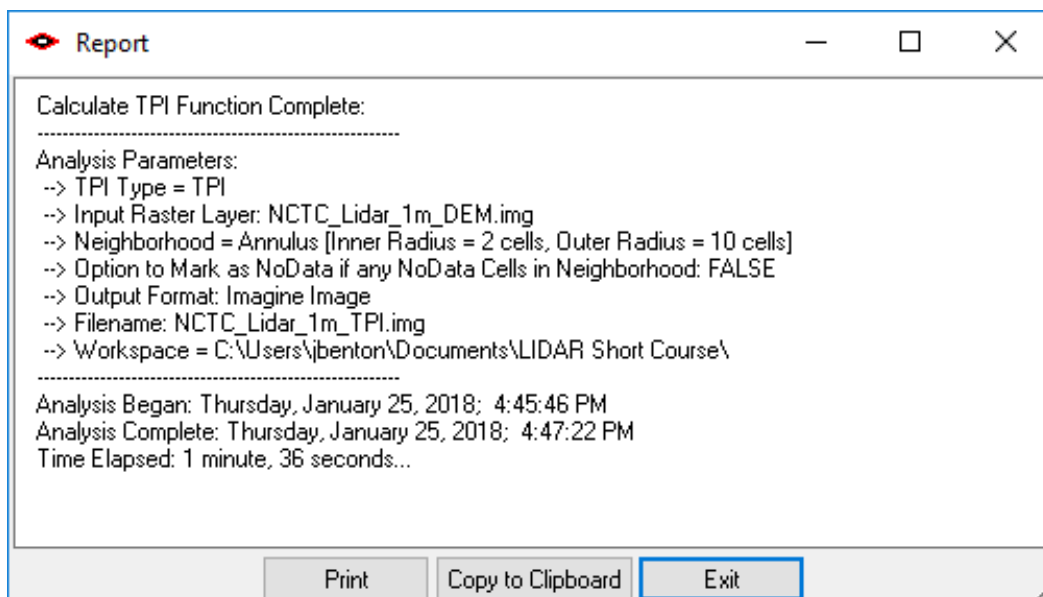


- Set the **Select DEM Raster Layer** to *NCTC_Lidar_1m_DEM.img* and **Select TPI Type** as *TPI*.
- Although there are many ways to calculate the TPI, we have found that an Annulus (donut-shaped) **Neighborhood Shape**, or window, with an **Inner Radius = 2 cells**, and **Outer Radius = 10 cells** works very well on a 1 m resolution DEM for identifying closed depressions. Set these options in the dialog box and set the **Output Raster Format** as *Imagine Image*.

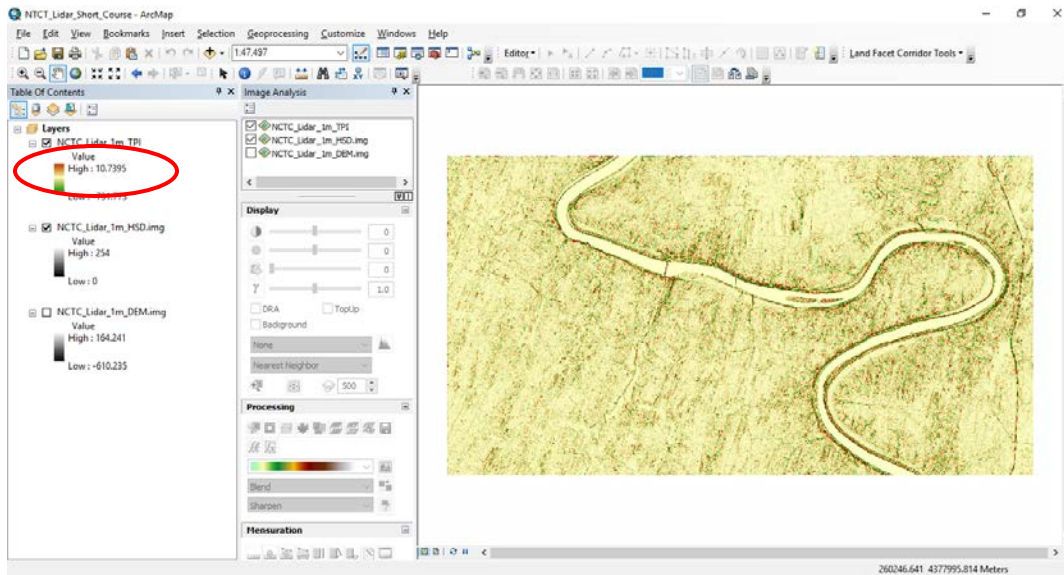
- Set the **Output Raster Dataset Name** to be NCTC_Lidar_1m_TPI.



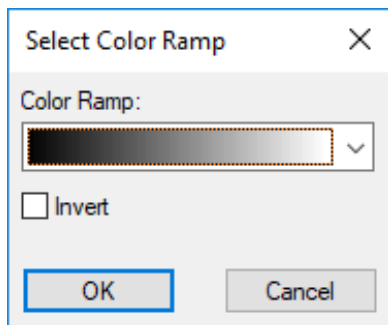
Click **OK**, and the TPI processing will begin. There is no way to monitor the progress of the TPI calculation, so please be patient. Depending on the processing speed of the computer, it could take anywhere from 1 to 5 minutes for this DEM. For a DEM the size of a typical 1:24,000 scale topographic quadrangle, the processing time could take between 10 minutes to half an hour. When completed, a dialog box should pop-up to inform you that the function is complete:



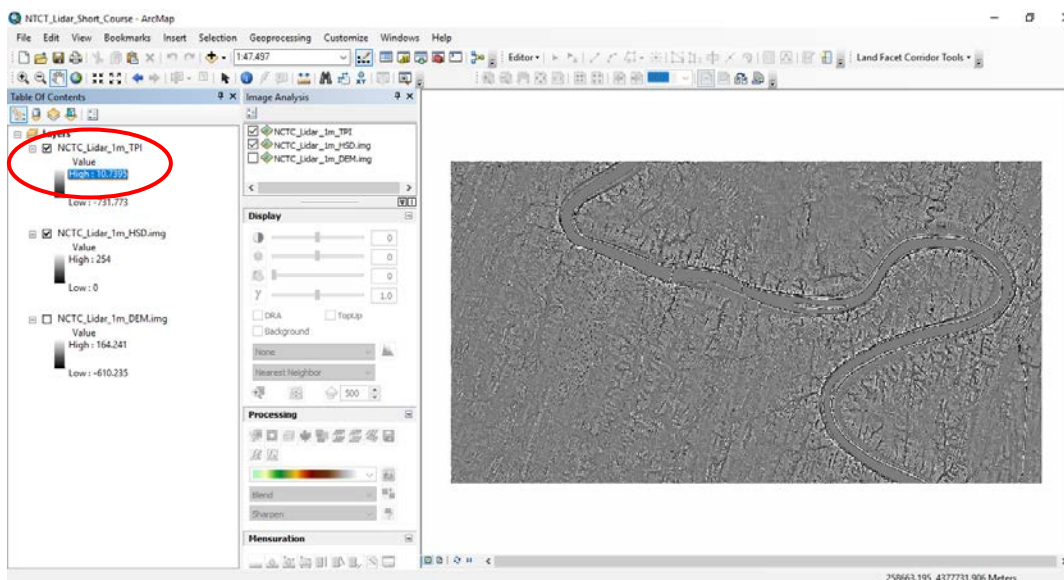
The resulting image should look something like what is below:



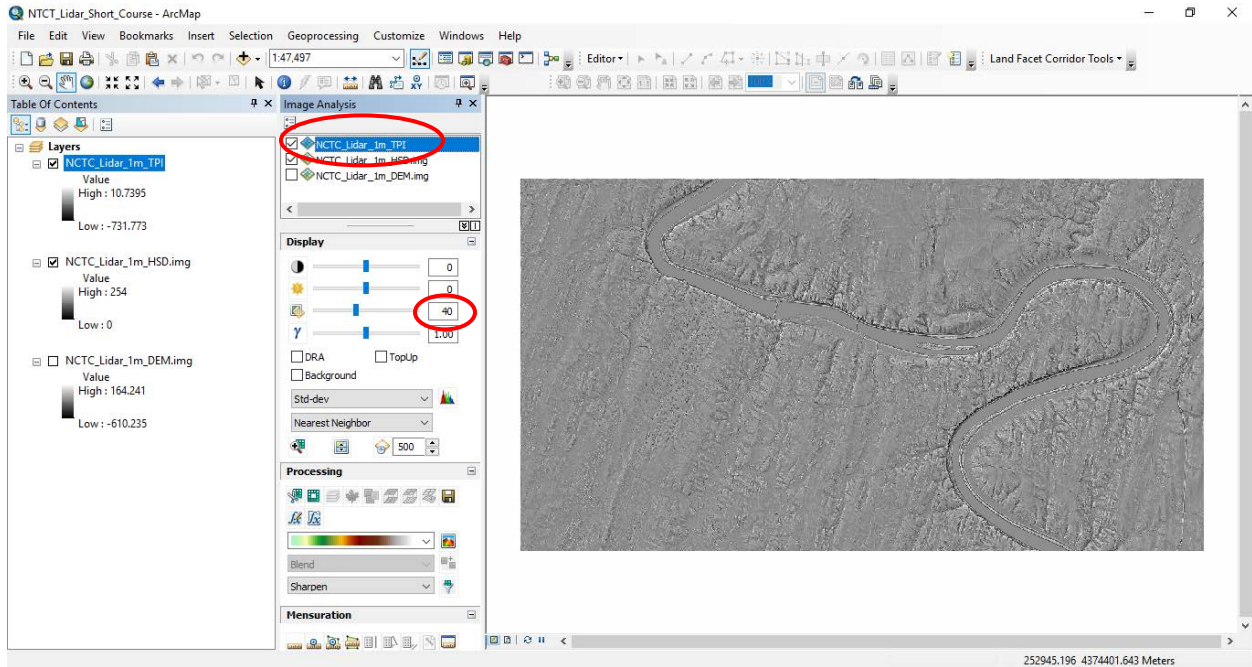
Left-click on the color ramp bar (encircled above), and change the color ramp to be black and white:



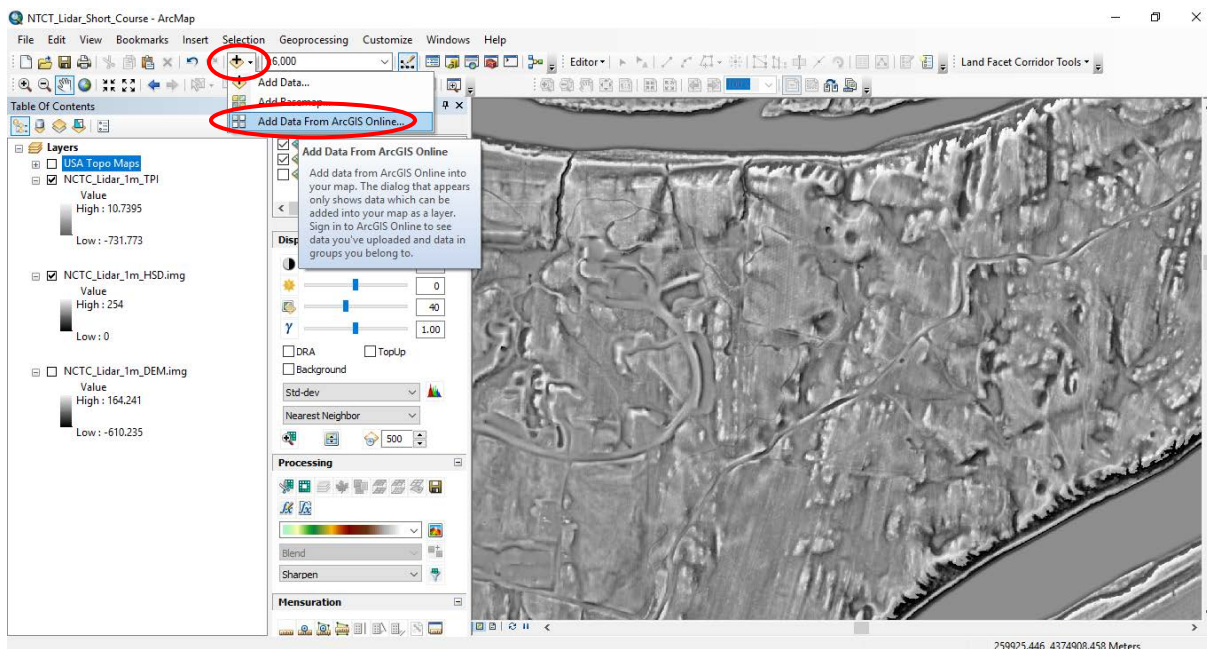
Your resulting TPI image should look like this:



Next, adjust the transparency of the TPI image, using the **Image Analysis** window, to 40% transparent so that the hillshade becomes visible beneath it. You can also do this by right-clicking on the TPI raster in the **Table of Contents** and going into the **Properties** → **Display** tab → **Transparency**:

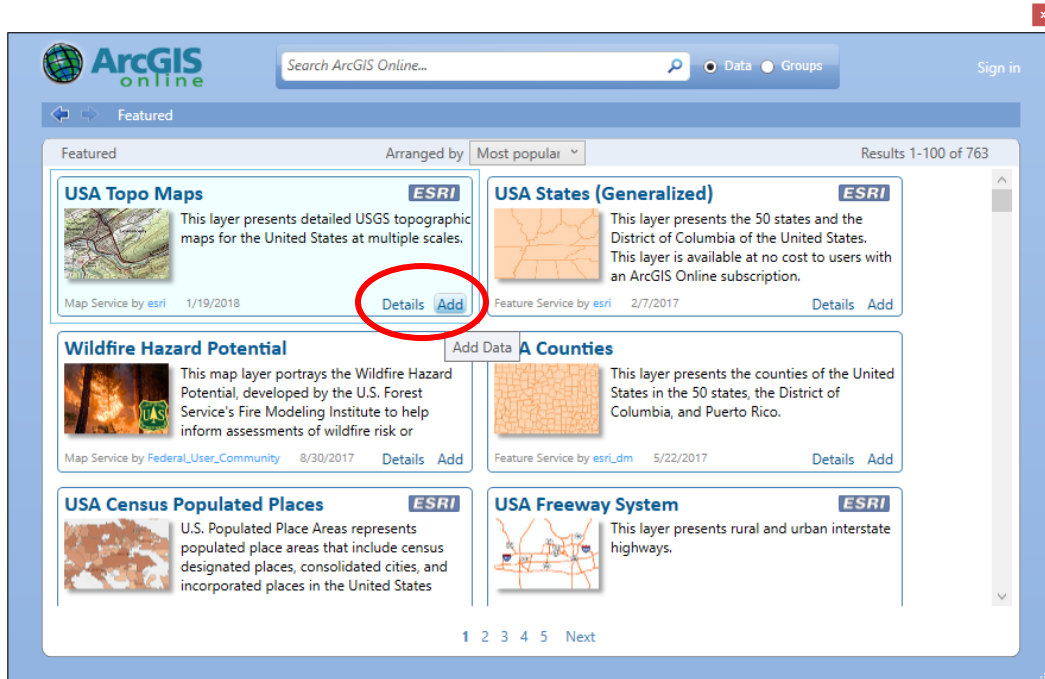


Zoom in to where the National Conservation Training Center is located, within the neck of the large meander bend in the Potomac River. Note that the lidar-derived DEM has cultural features removed such as houses and buildings. Often when buildings are removed from the elevation model, false closed depressions will result where the buildings once were located.

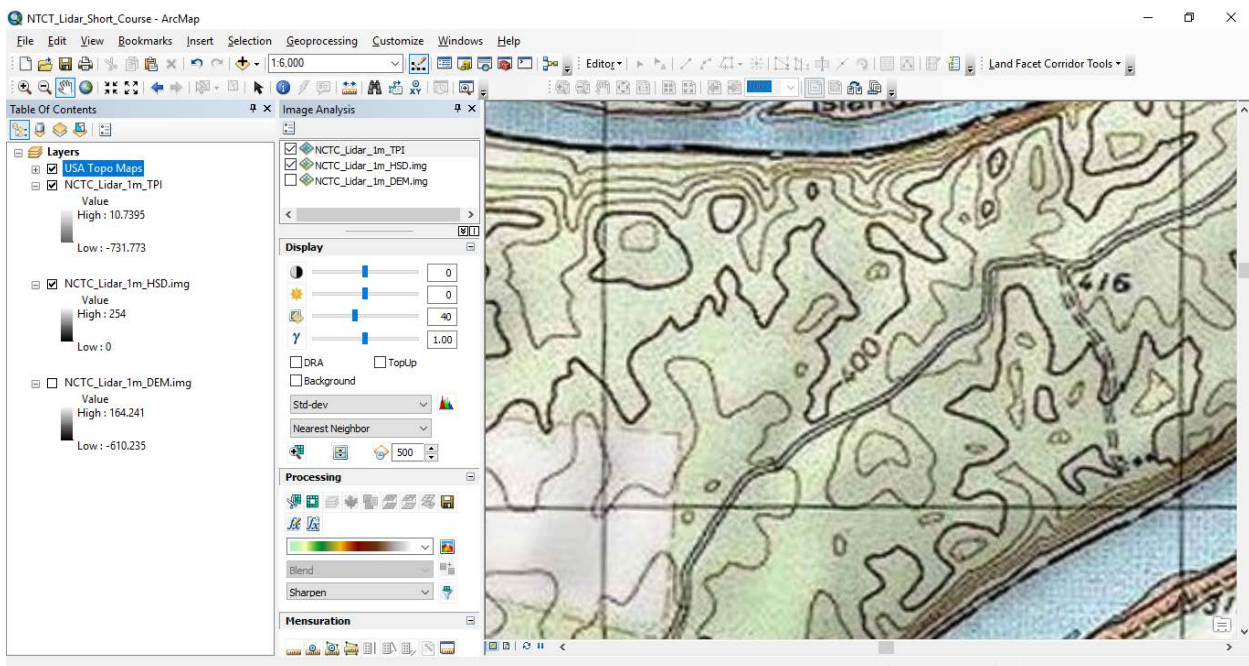


ADD TOPOGRAPHIC MAP & AERIAL IMAGERY OVERLAY FOR COMPARISON TO THE ELEVATION MODEL IMAGERY:

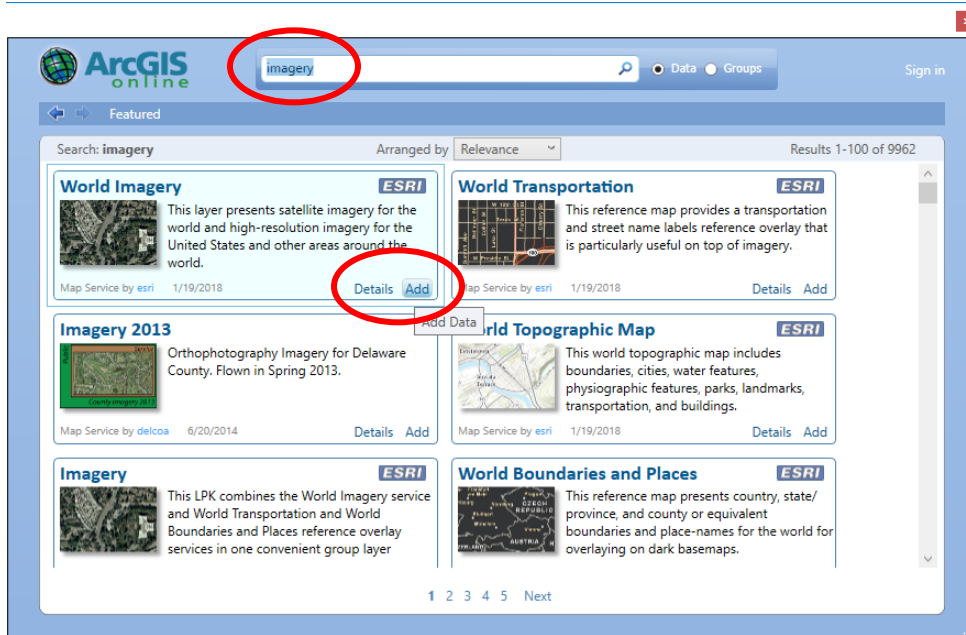
It is useful to compare the elevation imagery to a USGS topo map and actual aerial imagery for some context and reference. Go to **Add Data...** → **Add Data From ArcGIS Online...** and add *USA Topo Maps*:



You should see something like the image below. Note where the closed elevation contours are located, and toggle the imagery on and off or use the transparency slider to compare with the elevation imagery.

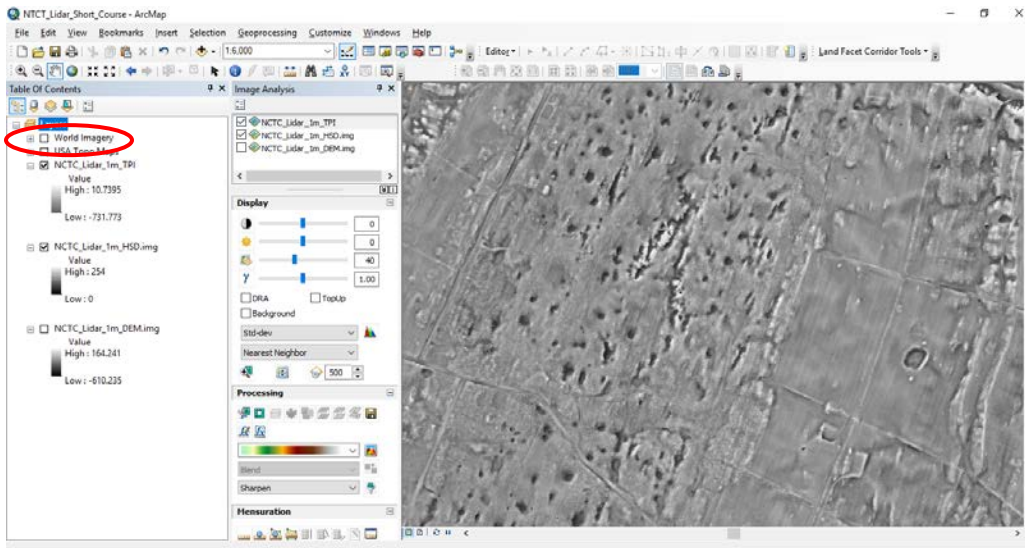


Now, return to **Add Data... → Add Data From ArcGIS Online...** and type “imagery” in the search bar, then add *World Imagery* when it appears in the list of available datasets:

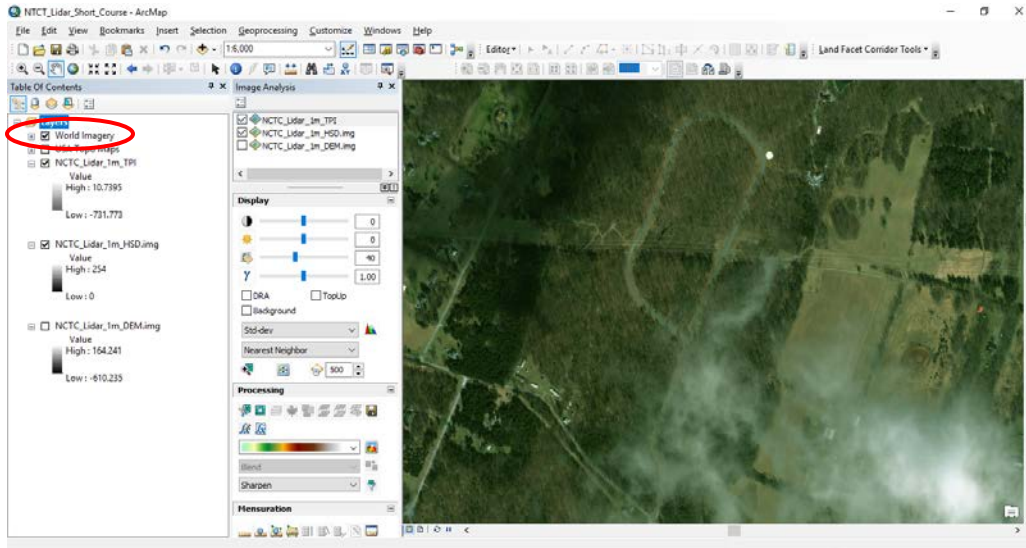


Toggle between the aerial imagery, the lidar imagery, and the USGS topographic map to compare. Adjust transparency among the different datasets.

Next, zoom to the area of high density closed depression along the south bank of the Potomac River, due west of the NCTC. This area is called Shepherd’s Cove, and the lidar was acquired prior to its development into lots for residential housing. This area will be visited on the Friday afternoon field trip. Compare the *World Imagery* to the DEM. Why might there be such a high density of closed depressions in such a localized area?



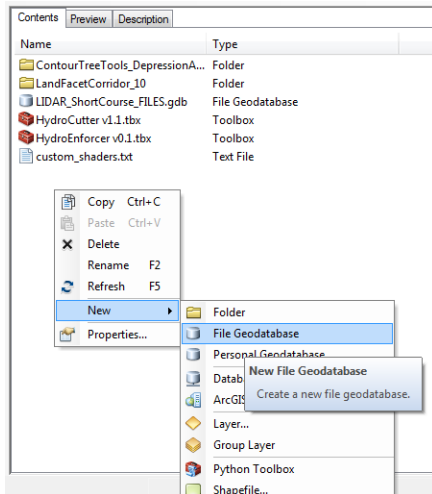
Note the road loop that exists in the aerial imagery below. This road was constructed after the lidar was flown, so it is not evident in the elevation model. Also note that much of the area with a high concentration of sinkholes is wooded in the aerial image, and was not likely to have been farmed, as opposed to the cleared area immediately to the west which has a much smoother appearance in the lidar imagery.



SEMI-AUTOMATED SINK DETECTION

The first step to working with the tools for semi-automated detection of closed depressions is to create an empty geodatabase within which to conduct the calculations, and to store all input as well as output files. Although a file geodatabase was already provided for you named *LIDAR_ShortCourse_FILES.gdb* immediately below we outline how we did this, so you can repeat this process after the course. **We will NOT need to work through the next two steps:**

1. To create a file geodatabase, open ArcCatalog, navigate to where you would like to store your database on your hard disk, right-click in the folder, and select **New** → **File Geodatabase**.



2. Rename the database following good naming conventions as outlined in the Note on Page 1., then right-click on the icon and select **Import** → **Feature class**. The first feature class you import will set the projection and datum of the database, so be sure it is properly projected to match the DEM and other data that will be in the database.

Again, for this course, a geodatabase *LIDAR_ShortCourse_FILES.gdb* has already been created for you as a workspace.

PART 1: Using the fill-difference method

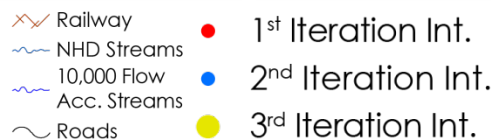
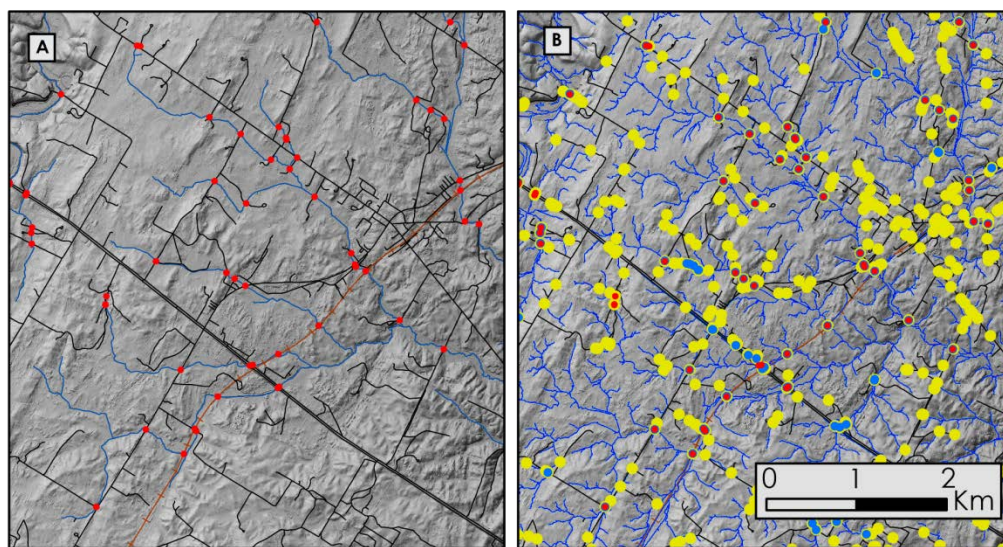
The fill-difference method finds closed depressions by “filling” all depressions in an elevation surface to their spill point according to a flow-routing algorithm, and subsequently subtracting the filled elevation surface from the original DEM in order to define the area and depth of all depressions within the area of interest. In order for the flow-routing algorithm to work properly, “digital dams” where streams flow under roads or other transportation routes must first be removed by “cutting” and “burning” new stream paths across the digital dams.

In this course, we will make use of several tools designed for use within ArcGIS to facilitate the mapping of closed depressions within a lidar-derived elevation model (Wall et al., 2015). A brief summary of the overall workflow is as follows:

- The *Cutter* tool allows streams to pass across roads and removes the "digital dams" by creating a new, modified “burned_DEM” with stream crossings burned across roads.

- The *Hydro* tool creates artificial stream lines that conform to the lidar topography, as well as a filled DEM for defining closed depressions. This tool defines stream lines in the lidar-derived elevation surface according to a user-defined flow accumulation value (i.e., the upstream contributing area in number of DEM pixels), and creates a filled DEM that has filled each closed depression to its spill point.
- The *Sinker* tool finds the closed depressions in the "burned_dem" created from the *Cutter* tool. It is best to run the *Sinker* tool once a DEM has been conditioned to the point where all known or suspected "digital dams" have been removed by iteratively running the *Cutter* and *Hydro* tools, with progressively lower flow accumulation values set in the *Hydro* tool. This allows the artificial stream network generated to increase in density upstream (see figure below), thus intersecting more roads, and improving the overall hydrologic conditioning of the DEM prior to extracting closed depressions.

NOTE: The *Hydro* tool will take the longest time to run, as it takes more processing power to generate the stream line vector data.



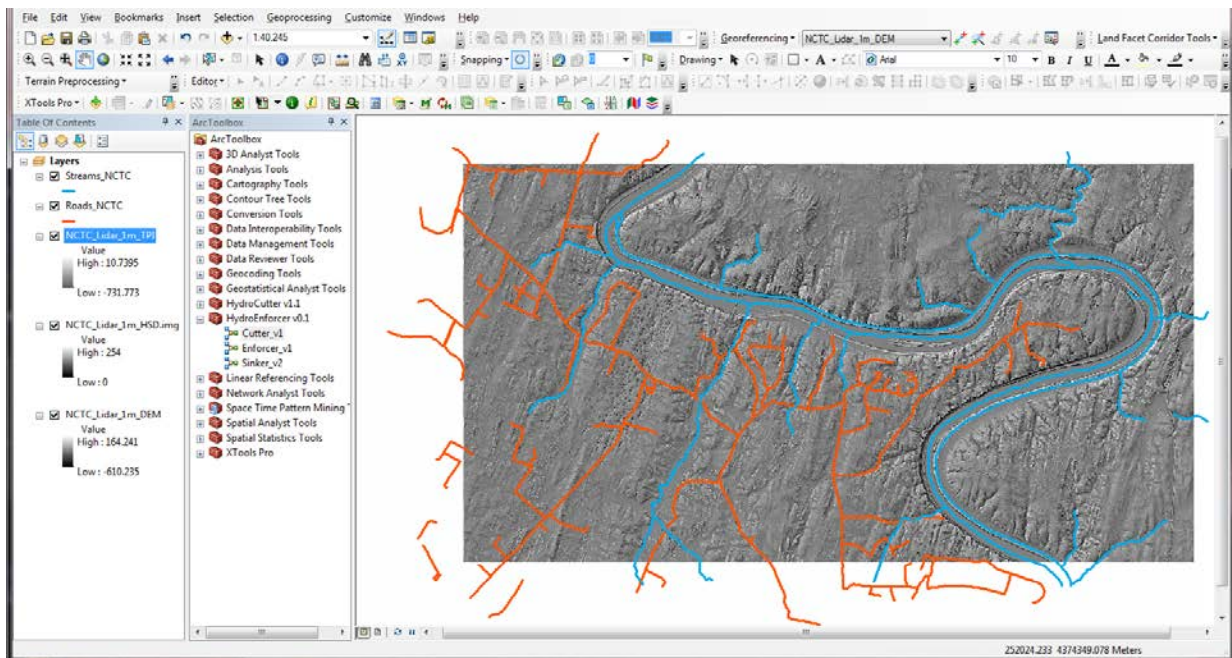
A typical workflow will be:

TOOL	INPUT	OUTPUT
1. Cutter	Roads (polylines)	burned_DEM (raster)
2. Hydro	Streams from the NHD (polylines) burned_DEM (raster) ←	streams_400k (polylines) Flow Accumulation @ 400,000
3. Cutter	Roads (polylines) streams_400k (polylines) ←	burn_DEM_400k (raster)
4. Hydro	burn_DEM_400k (raster) ← Flow Accumulation @ 10,000	streams_10k (polylines)
5. Cutter	Roads (polylines) streams_10k (polylines) ←	burn_DEM_10k (raster)
6. Hydro	burn_DEM_10k (raster) ← Flow Accumulation @ 2,000	streams_2k (polylines)

It is recommended to run through the *Cutter* and *Hydro* tools at least 2 additional times after the first run with the NHD streams, since the new stream lines created by the *Hydro* tool will better conform to the lidar-derived topography. In addition, creating stream lines defined at progressively lower flow accumulation thresholds will extend these lines further upstream, and therefore will intersect more road crossings after each running of the *Hydro* tool.

We recommend running the *Hydro* tool at flow accumulation thresholds of 400,000, then 10,000 after first running it using the National Hydrologic Dataset streams on 1 m resolution DEMs. However, for this exercise, we will only work with the 10,000 flow accumulation stream lines for demonstration purposes.

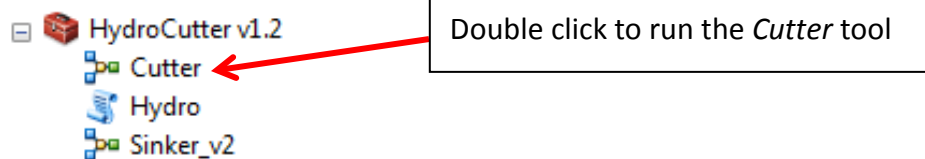
Within the ArcMap project that was begun above, go to **File** → **Add Data**, and load the polyline vector datasets *Roads_NCTC* and *Streams_NCTC* that are in the *LIDAR_ShortCourse_FILES.gdb* geodatabase to the ArcMap project. These will now be shown in the **Table of Contents**.



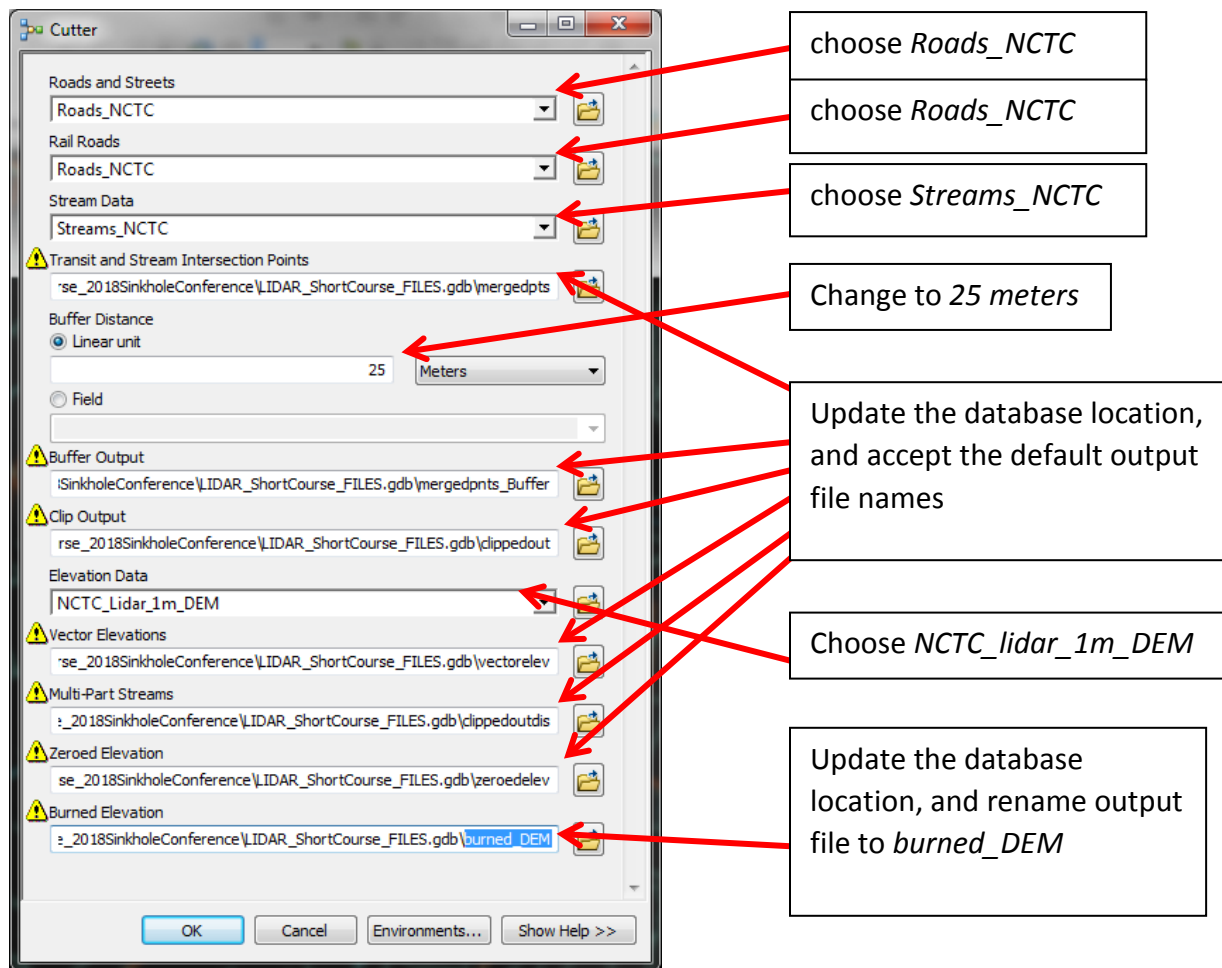
Run the Cutter tool

The *Cutter* tool is used for removing “digital dams” where streams cross roads. The tool finds the intersection between user-supplied road and stream vectors, and then cuts a path through the elevation model through the roads to permit flow across those artificial barriers where culverts exist in reality.

Open the *HydroCutter v1.2* toolbox. There are three toolsets inside:



Run *Cutter* first by double clicking on the tool in the toolbox. This tool is designed to breach “digital dams” that may occur within the DEM at road and stream intersections, based on the road and stream vector polyline data that are supplied by the user.



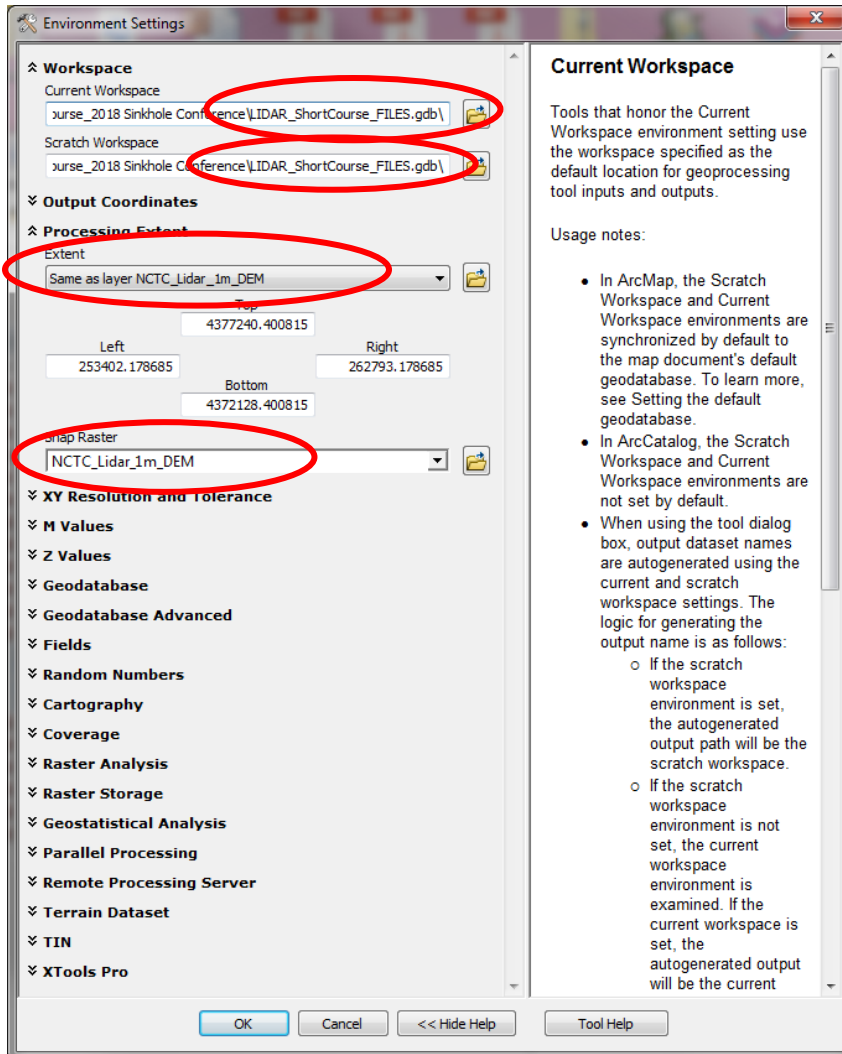
- Set the **Roads** and **Streets** dialog box to *Roads_NCTC*, and the **Rail Roads** to the same *Roads_NCTC* shapefile, as there are no railroads in this DEM area.
- Set the **Stream Data** dialog box to *Streams_NCTC*.
- Change all output locations to be within *LIDAR_ShortCourse_FILES.gdb* which is in the folder on the desktop. This can be done by clicking the **Navigate to Data** button (folder icon) and navigate to the file geodatabase.
- Once you are in the file geodatabase input a name for the output dataset. You can then highlight the base folder pathname and press **CRTL + C** to copy the path. In the remaining outputs, highlight the incorrect pathname and paste the correct name pressing **CRTL + V**.

NOTE: Be sure to remove all the *.shp* extensions from the file names, as they will be placed within the geodatabase instead of written to standalone shapefiles (if not done, this can result in the tool crashing with a “Table Name is Invalid” error). An example of the proper pathname would be:

C:\Users\username\Desktop\LidarShortCourse_2018SinkholeConference\LIDAR_ShortCourse_FILES.gdb\mergedpnts

- Set **Buffer Distance Linear unit** to *25 meters*.
- For the **Elevation Data** choose *NCTC_Lidar_1m_DEM*.

- Be sure to place your **Burned Elevation** data into the file geodatabase and we would recommend naming it *burned_DEM*. This will be the hydrologically conditioned DEM surface that will be used to find the closed depressions or in other iterations of the *HydroCutter* toolbox.
- Next, click on the **Environments...** button.
- Change the settings of the **Current Workspace** and **Scratch Workspace** to be *LIDAR_ShortCourse_FILES.gdb*.
- Change the **Processing Extent** to be the same as the *NCTC_Lidar_1m_DEM*, and set the **Snap Raster** to *NCTC_Lidar_1m_DEM*.



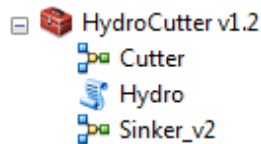
Click **OK**, and **OK** again to run the tool. When it has finished, you should see several additional layers added to the **Table of Contents**: *mergedpts*, *clippedoutdis*, *clippedout*, *mergedpts_buff*, *burned_dem*, *zeroelev*, and *vectorelev*.

Turn off all but the *burned_dem* layer. This will be the input into the next step.

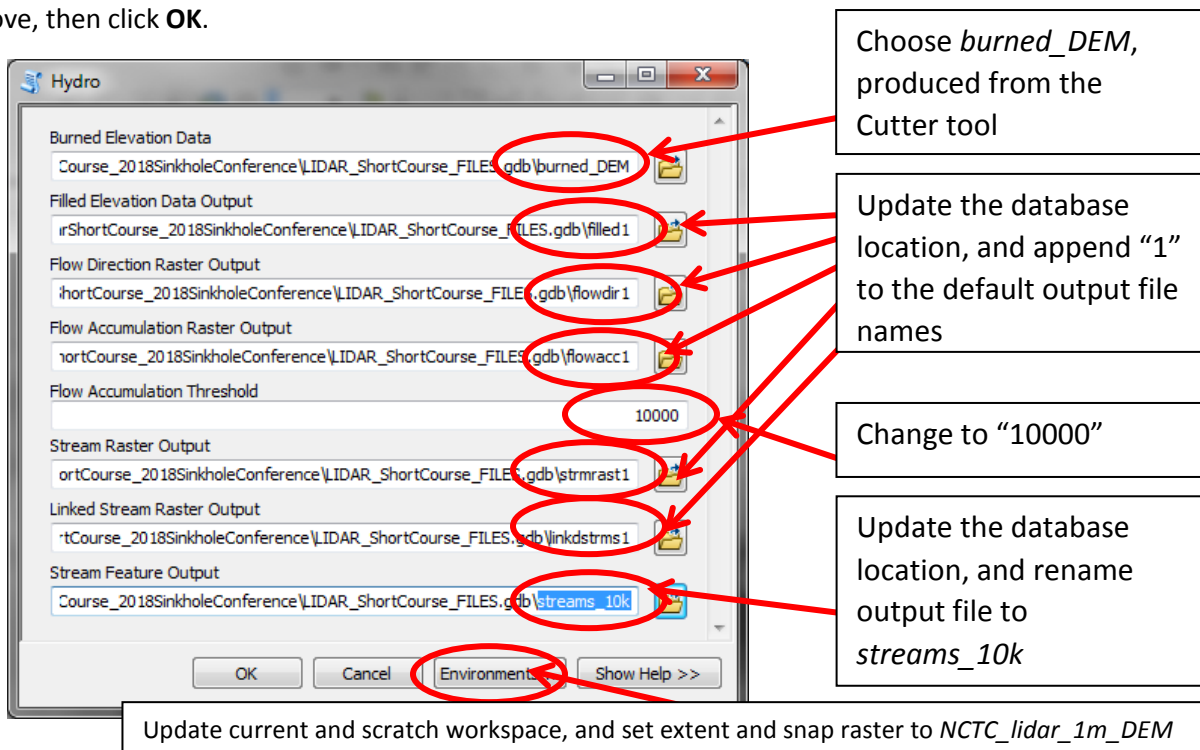
Run the Hydro tool to generate stream lines:

The *Hydro* tool is located within the *HydroCutter v1.2* toolbox. This tool is used to define streamlines within the lidar data according to a pre-defined flow accumulation threshold set by the user. The resulting streamlines can then be used to further filter out artificial depressions created at road/stream intersections not captured in the original user-supplied stream polylines obtained from the USGS National Map website, lines generated from a previous run of the Hydro tool, or other source.

Double click on the *Hydro* tool in the *HydroCutter v1.2* toolbox:

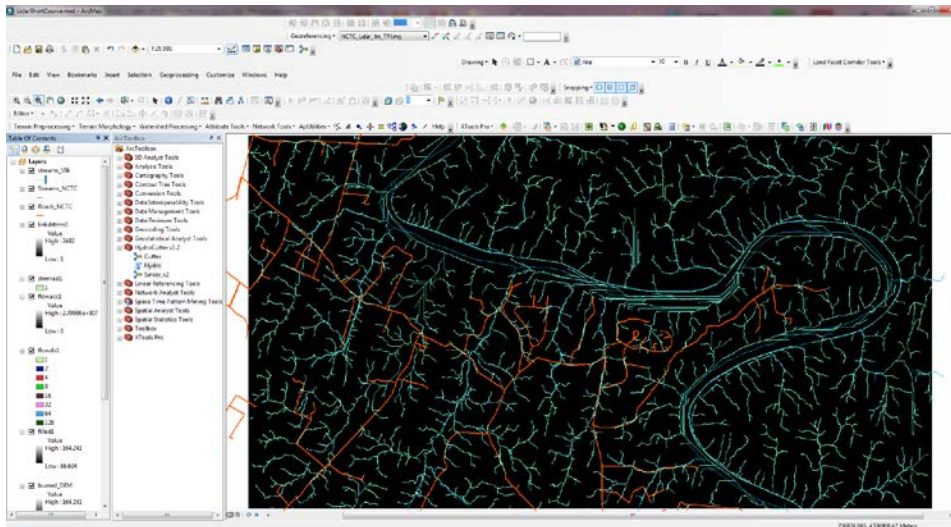


Select the *burned_dem* as the **Burned Elevation Data** input. Change all paths to be set to the **LIDAR_ShortCourse_FILES.gdb** just like what was done for the *Cutter* tool above. Append the number "1" to the end of all output file names. This will help keep the output organized later when the process is repeated. Set the **Flow Accumulation Threshold** to *10000*. Change the filename of the **Stream Feature Output** output to *Streams_10k*. Set the **Environments...** settings to those used for the *Cutter* tool as above, then click **OK**.



This tool could from **40 minutes up to an hour** to complete running.

When the tool execution is complete, you should see something similar to the image below:

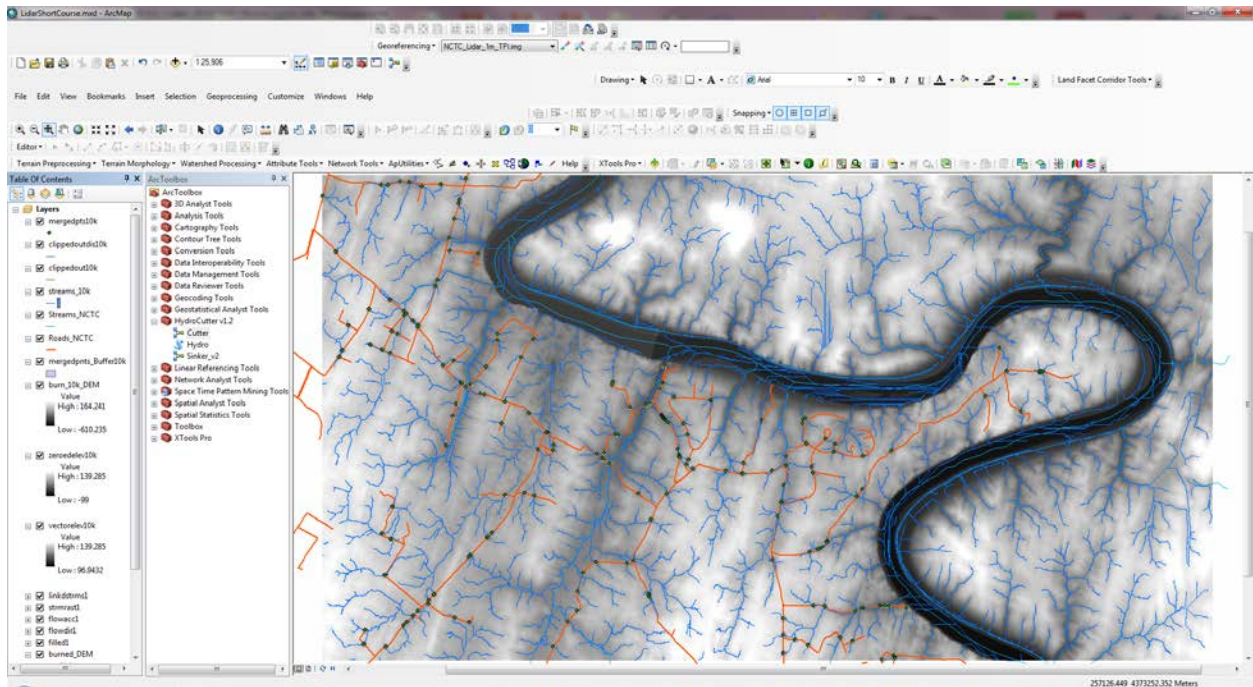


Next, run the *Cutter* tool once more to intersect the roads with the 10,000 stream lines generated from the *Hydro* tool. Use the same settings as before, except set the **Stream Data** input to *streams_10k*.

The screenshot shows the 'Cutter' tool dialog box with several annotations. The 'Stream Data' input is set to 'streams_10k'. The 'Buffer Distance' is set to '25 Meters'. The 'Elevation Data' is set to 'NCTC_Lidar_1m_DEM'. The 'Burned Elevation' output is set to 'burn_10k_DEM'. The 'Environments...' button is circled in red.

- choose *streams_10k*, produced from the *Hydro* tool
- Change to 25 meters
- Update the database location, and append "10k" to the default output file names
- Choose *NCTC_lidar_1m_DEM*
- Update the database location, and rename output file to *burn_10k_DEM*
- Update current and scratch workspace, and set extent and snap raster to *NCTC_lidar_1m_DEM*

The result should look something like the image below, with many more road/stream intersection points shown than before



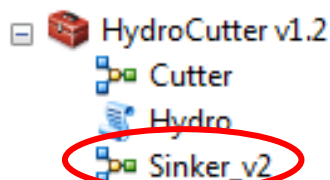
The resulting *burn_10k_DEM* will be final output raster, conditioned to remove “digital dams” at these road/stream intersections.

We are now ready to identify the closed depressions within this DEM!

Run the Sinker tool

In the *Sinker* tool, will use the burned DEM created in the previous step (*burn_10k_dem*) which will be hydrologically conditioned using the *Fill* geoprocessing algorithm in ArcMap (found at **ArcToolbox** → **Spatial Analyst Tools** → **Hydrology**). The original DEM is then subtracted from the filled DEM resulting in a depth raster. The depth raster is then thresholded at a user defined value (a minimum of two times the vertical RMSE is recommended). This threshold raster is then converted to a polygon dataset. An area threshold is then applied to the sink polygon dataset.

Double click on the *Sinker_v2* tool in the *HydroCutter v1.2* toolbox:



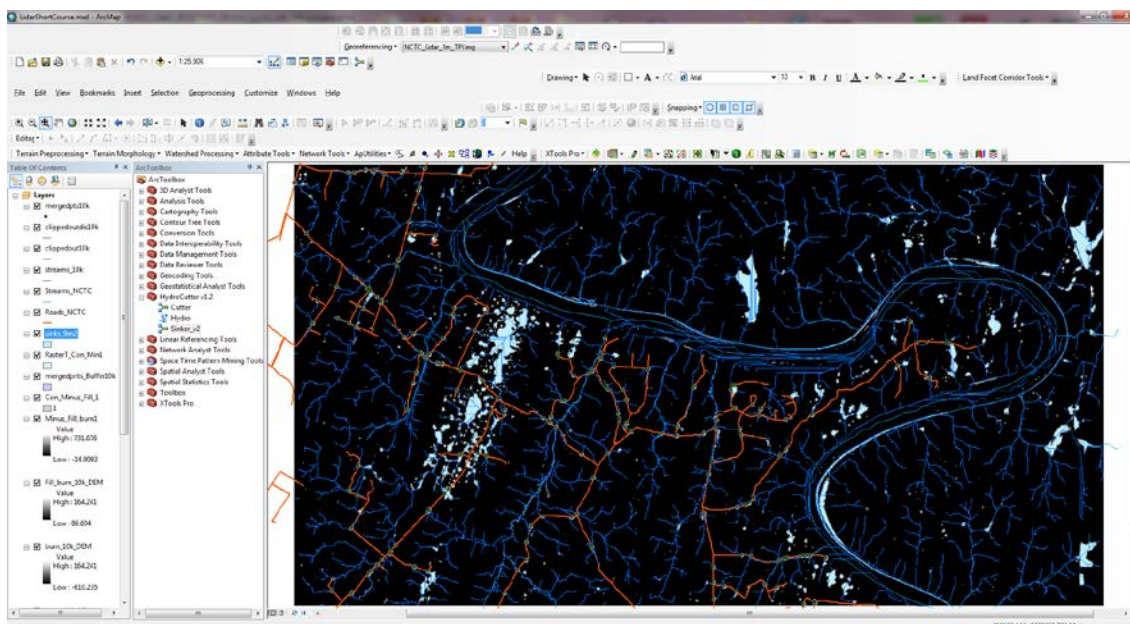
Set the **Expression VALUE** > 0.2, and the **Area Threshold for Sinkholes** to *Shape_Area* > 9. This sets a minimum depth of 0.2 m and a minimum area of 9 square meters to the resulting sinkhole data that are created.

Change all paths to be output data to *LIDAR_ShortCourse_FILES.gdb* just as was done before. Additionally, remove all file extensions to the output as before, and rename the output of the **Threshold Sinkholes** to be *sinks_9m2*:

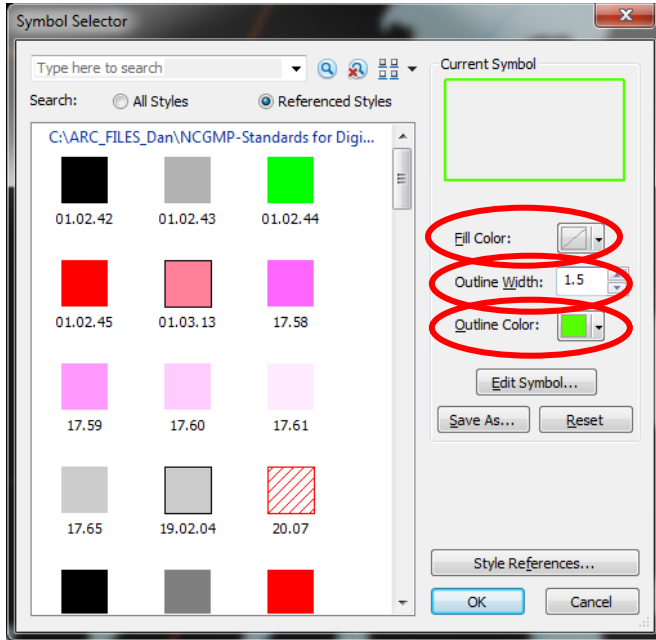
The screenshot shows the **Sinkers_v2** tool dialog box with the following settings and callouts:

- Original DEM:** *NCTC_Lidar_1m_DEM* (Callout: Choose *NCTC_Lidar_1m_DEM*)
- Burned DEM:** *burn_10k_DEM* (Callout: Choose the *burn_10k_DEM* output from Cutter)
- Filled DEM:** *:\reference\IDAR_ShortCourse_FILES.gdb\Fill_burn_10k_DEM* (Callout: Rename output to *Fill_burn_10k_DEM*)
- Depth Raster:** *:\reference\IDAR_ShortCourse_FILES.gdb\Minus_Fill_burn_10k_DEM* (Callout: Update output geodatabase)
- Expression (optional):** *VALUE > 0.2* (Callout: Set minimum sinkhole depth to 0.2 m)
- Sink Raster:** *:\reference\IDAR_ShortCourse_FILES.gdb\Con_Minus_Fill_burn_10k_DEM* (Callout: Update output geodatabase)
- Sink Polygons:** *:\reference\IDAR_ShortCourse_FILES.gdb\RasterT_Con_Minus_Fill_burn_10k_DEM* (Callout: Update output geodatabase)
- Threshold Sinkholes?**
- Area Threshold for Sinkholes (optional):** *Shape_Area > 9* (Callout: Set minimum sinkhole area to 9 m²)
- Threshold Sinkholes:** *:\reference\IDAR_ShortCourse_FILES.gdb\sinks_9m2* (Callout: Update current and scratch workspace, and set extent and snap raster to *NCTC_lidar_1m_DEM*)

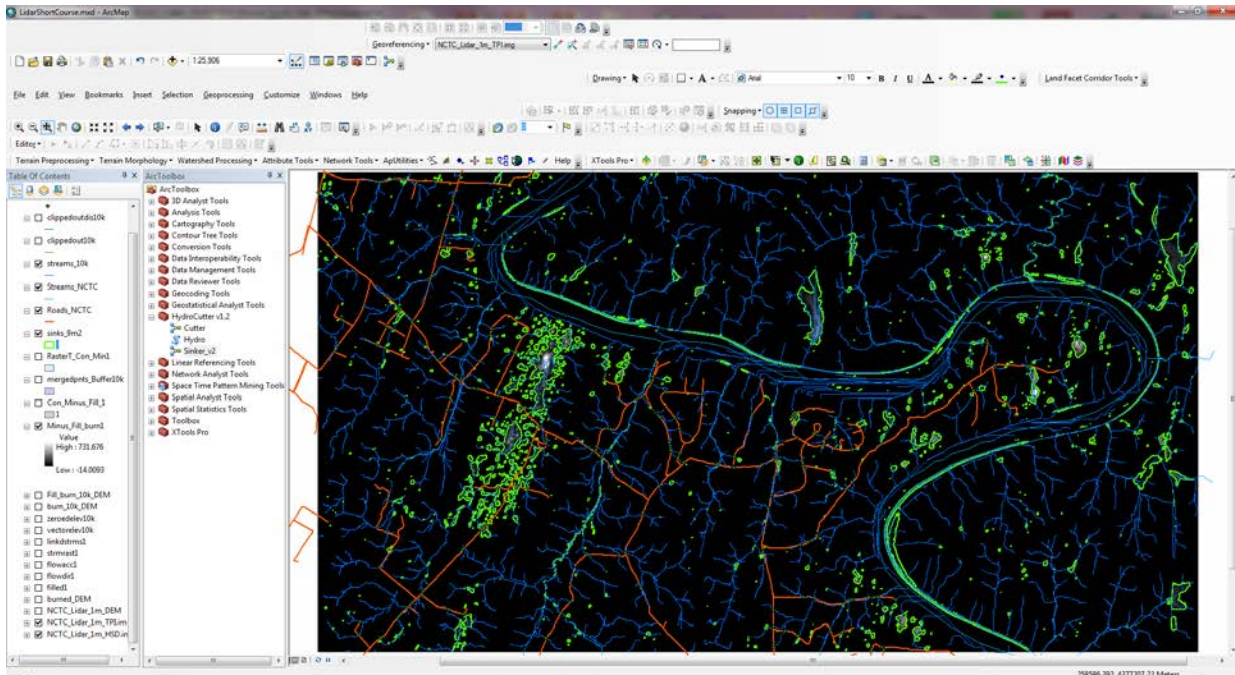
Set the **Environments...** settings to those used for the *Cutter* tool as above, then click **OK**. When the tool execution is complete, you should see something similar to the image below:



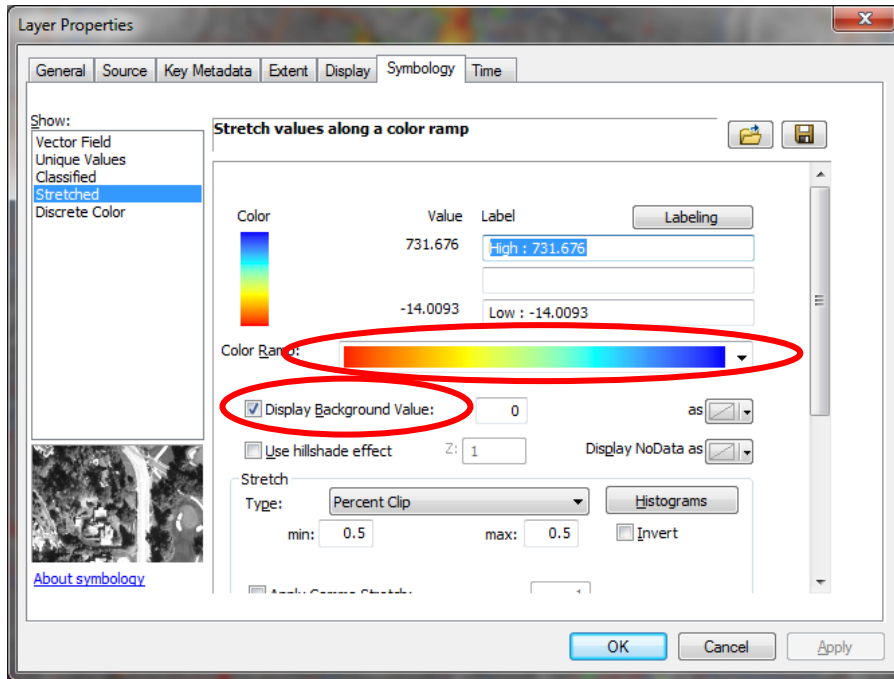
In the **Table of Contents**, uncheck all layers except *sinks_9m2* and *Minus_Fill_burn1*. Click on the polygon symbol of the *sinks_9m2* layer to open the **Symbol Selector** dialog, click on the button next to **Fill Color** and click **No Color**. Set the **Outline Width** to 1.5 and the **Outline Color** to *Medium Apple* (bright green).



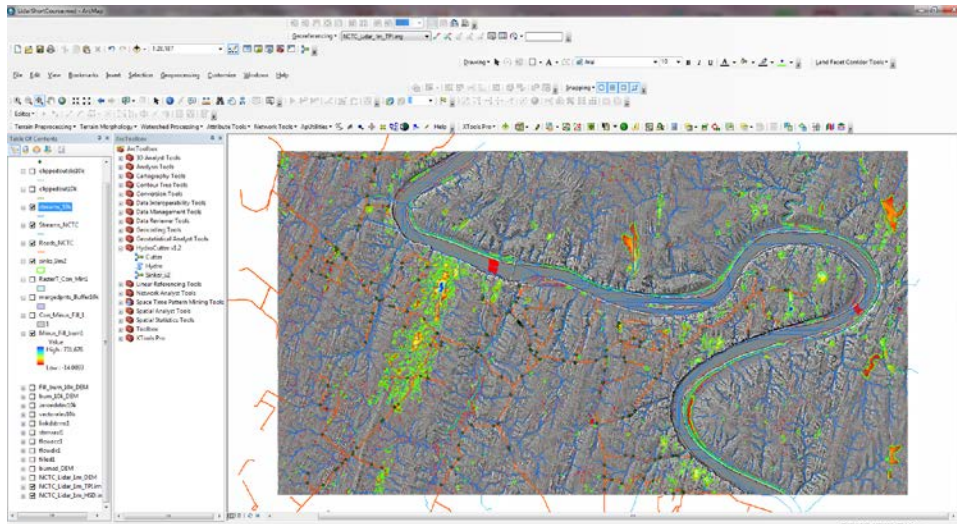
You should see something like the image below:



Next, right-click on the *Minus_Fill_burn1* raster layer, and select **Properties** → **Symbology**. Change the **Color Ramp** to be *Spectrum-Full Bright* (from blue through yellow to red), and check the box to **Display Background Value 0** as *No Color*:



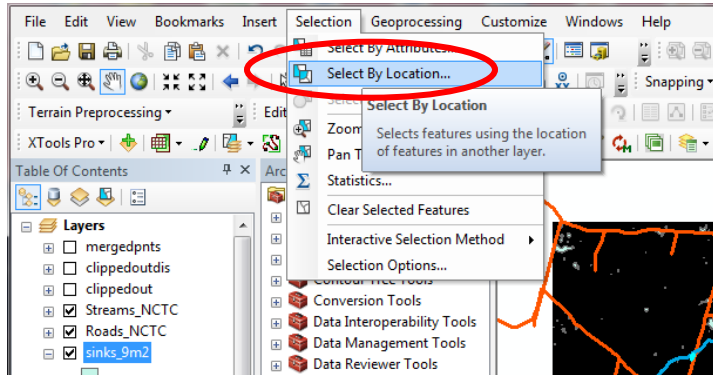
Now you should see something like the image below. The background value of 0 depicts all pixels that have not changed from the original DEM, since the *Minus_Fill_burn1* DEM is the subtraction of the original DEM from the filled burned DEM; the color ramp will highlight the depths of the depressions that occur:



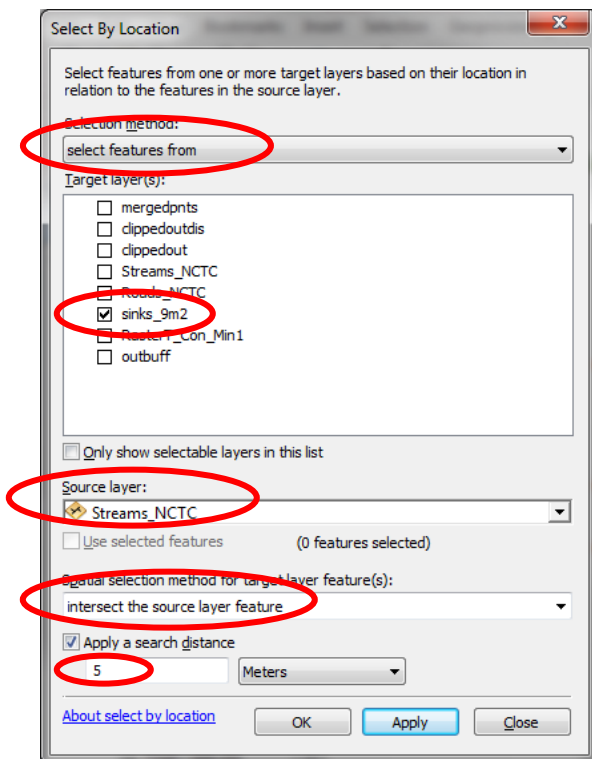
You now have a polygon layer that is representative of the possible sinks from the difference raster based upon the accuracy of the lidar data. Zoom in and explore the results. Where is the deepest depression? Why might depressions show up in the river channel (hint: they are not real). Why are there small spec-like depressions in the fill-difference DEM (*Minus_Fill_burn1*) that are not outlined in green like the *sinks_9m2* polygons?

Clean up the depression data:

Remove spurious sinks along stream courses by first selecting all sinks that intersect a stream within a 5 meter buffer width along the stream line. Right-click on the *sinks_9m2* layer, and select **Edit Features** → **Start Editing**. Go to **Selection** → **Select by Location**



Under **Selection method**: choose *select features from*. Under **Target layer(s)** check the box next to the *sinks_9m2*. Set the **Source Layer** to *Streams_NCTC*. Set the **Spatial selection method for target layer feature(s)**: to *intersect the source layer feature*, and check the box next to **Apply a search distance** entering a value of *5 meters*.



Once the sinks that intersect the *Streams_NCTC* lines are selected, you may delete them by going to the top menu **Edit** → **Delete**.

It is also good to remove any artificial depressions (pits) that may have been introduced into the lidar elevation model caused by removal of buildings during the processing of the DEM (see **Appendix 2** for a

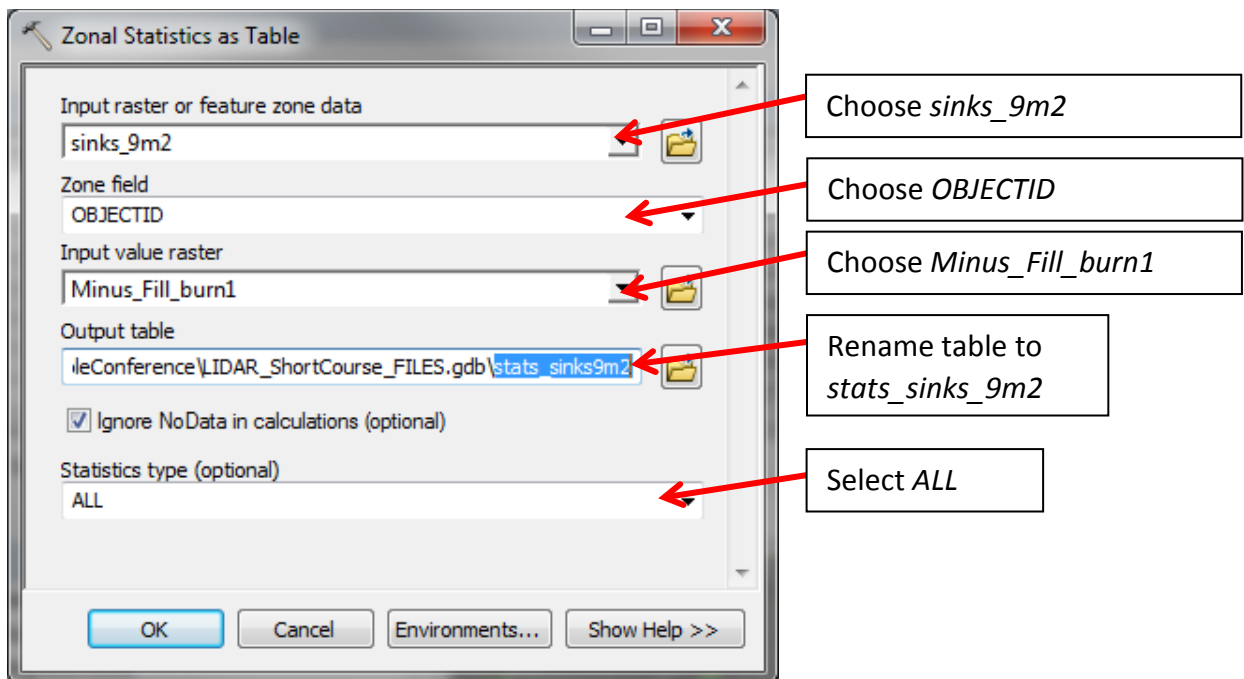
procedure to screen out erroneous pixels of negative elevation). For this, it is helpful to have a polygon layer of buildings that you may use to select sinks within a given buffer distance (e.g., 2 m) of a structure and delete them similarly to how sinks were removed from stream channels. Another means is to manually delete the artificial sinks using aerial photograph overlays to check for structures. We will not do this step here.

Additional QA/QC is needed to finalize the closed depressions inventory, and this is where the process ceases to be automated. It is advised to move across your area of interest using an index grid of ~1.5km x1.5km (the lidar tile index is a good grid for this purpose), and check the validity of each closed depression based upon 1) evidence in the lidar imagery (e.g. hillshade, TPI, etc.), 2) evidence in aerial imagery, 3) evidence from field checking. There is no substitute for going through your data with a fine-toothed comb, so to speak, in order to establish its validity. Note-taking and additions to the attributes of the sinks polygons along the way is also very useful, especially for depressions that may be ephemerally ponded, or may be stream sinking points, or the source of a spring, etc.

Calculate geometric characteristics of each sinkhole polygon:

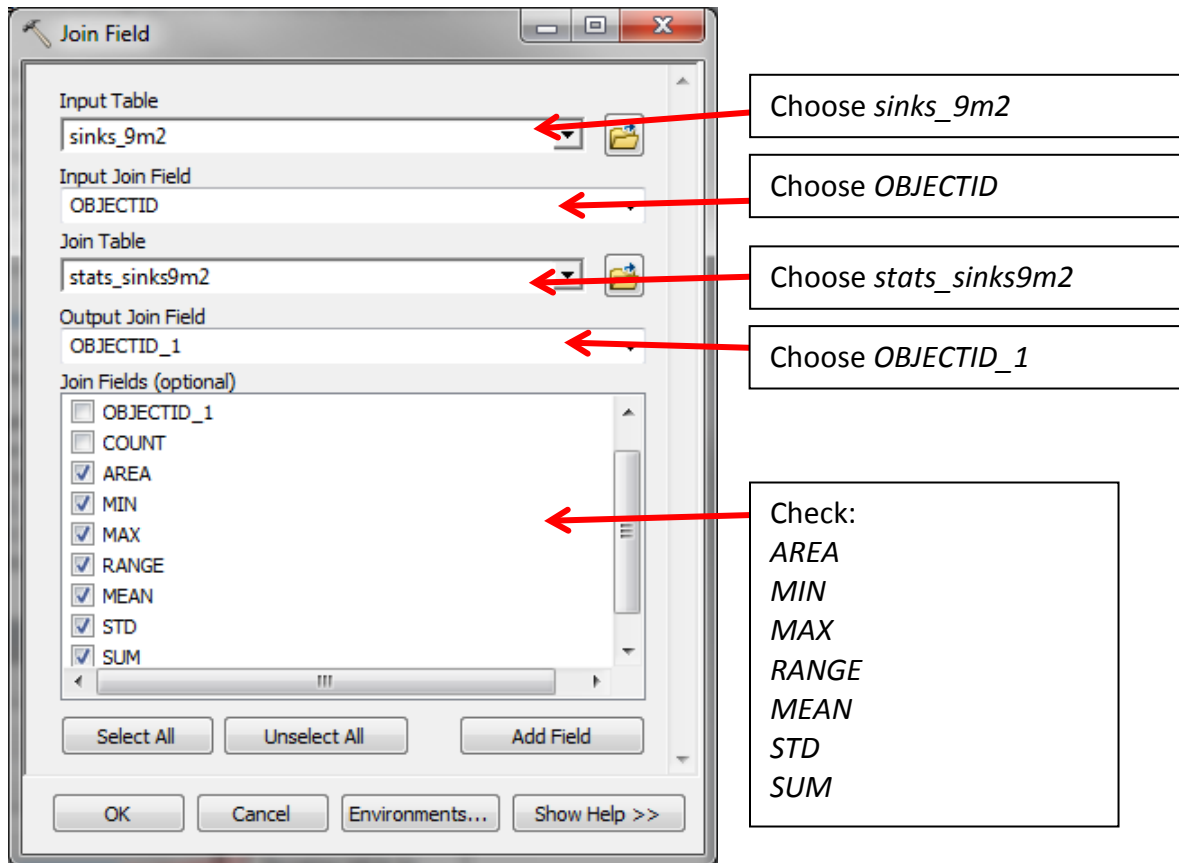
The steps below use the *Zonal Statistics* and *Zonal Geometry* tools in ArcMap to calculate some morphometric parameters of each sinkhole.

In **ArcToolbox** navigate to **Spatial Analyst Tools** → **Zonal** and open the *Zonal Statistics as Table* tool. We will use this tool to act on the raster *Minus_fill_burn1* (i.e. the fill-difference DEM or depth raster) to calculate the planimetric area (AREA), minimum depth (MIN), maximum depth (MAX), mean depth (MEAN), depth range (RANGE), standard deviation (STD) and sum (SUM) of the depths for all 1m pixels in each polygon in the *sinks_9m2* polygon dataset. Since the SUM is of the total added depths of each 1m pixel, the result equals the volume of the depression. If the pixel area was greater than 1m, one would multiply the integer value of the pixel area by the SUM to obtain the volume in cubic meters.



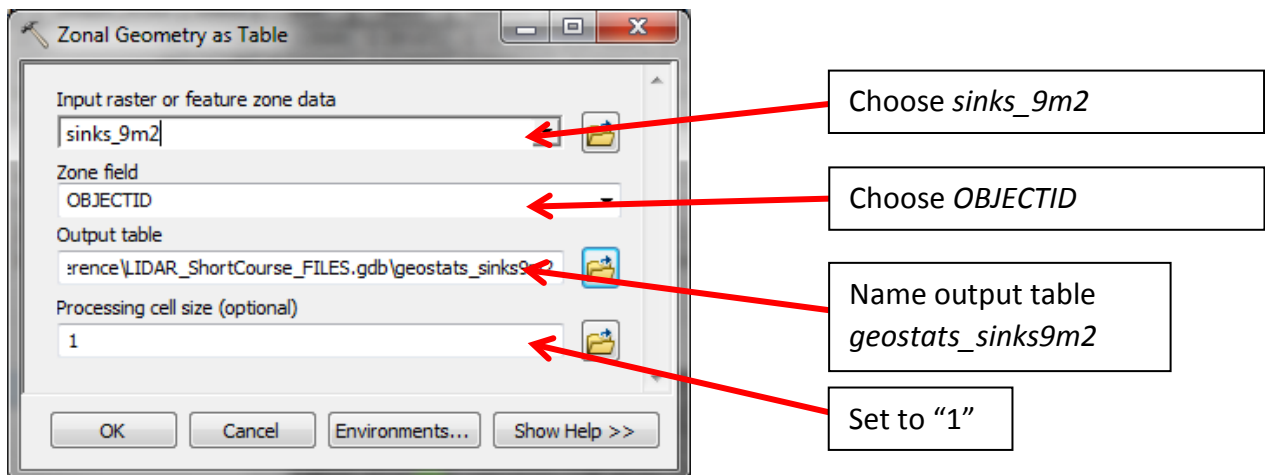
The results are written to a table that can then be joined to the attributes of the polygons. To join this table to the polygons, stop editing the *sinks_9m2* layer by going to **Editor** → **Stop Editing**.

Next, in **ArcToolbox** navigate to **Data Management Tools** → **Joins** and open the *Join Field* tool. Under **Input Table** input *stats_sinks9m2* for the table:

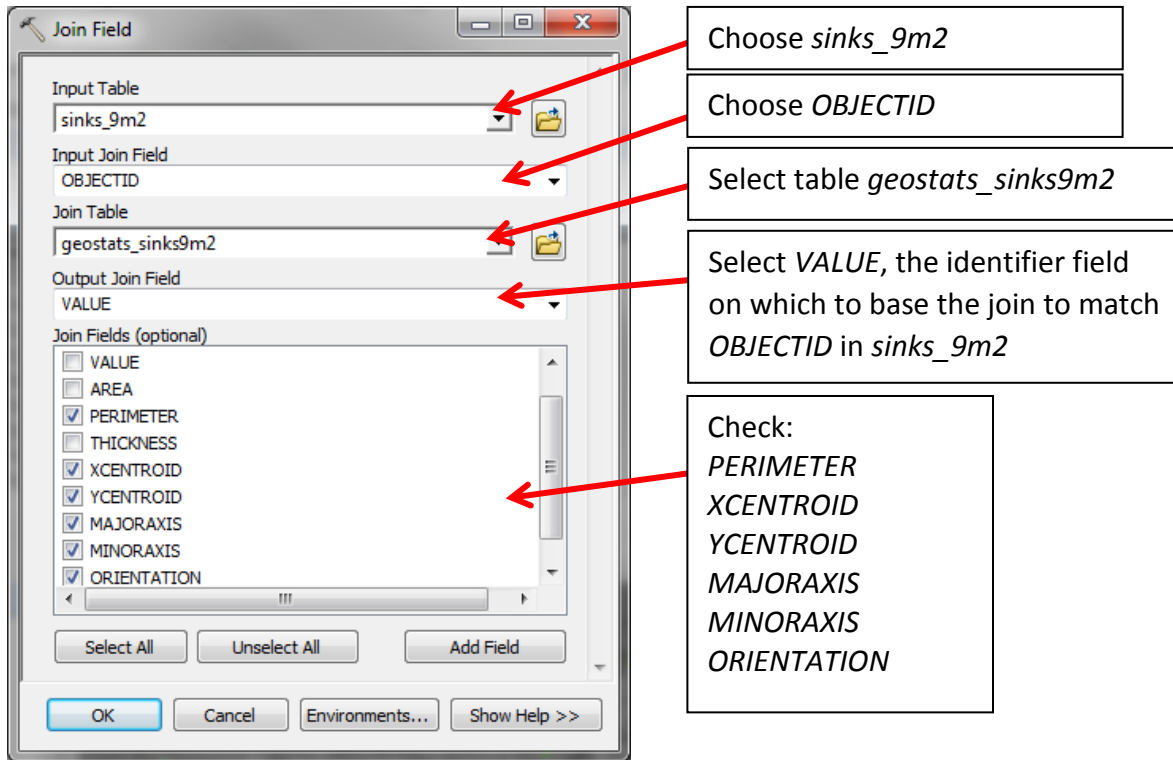


When the *Join Field* tool is complete, right-click on the *sinks_9m2* and click **Open Attribute Table** to see the results in the table. Note that these statistics are for calculated on the depths of each depression area.

Use the *Zonal Geometry as Table* tool to write the output geometry of the *sinks_9m2* polygons.



The *Zonal Geometry as Table* tool could take up to 10 minutes to run. The table will contain values of area, perimeter, centroid XY coordinates, and length of major and minor axes of the feature envelope ellipse. Navigate to the tool by returning to **ArcToolbox** → **Spatial Analyst Tools** → **Zonal** and click on *Zonal Geometry as Table*. Set the **Zone field** to *OBJECTID* and **Processing Cell Size** value to “1” to mimic the operation working on a raster of 1 m cell size. Use the *Join Field* tool as before to append the results to the *sinks_9m2* polygon feature attributes.



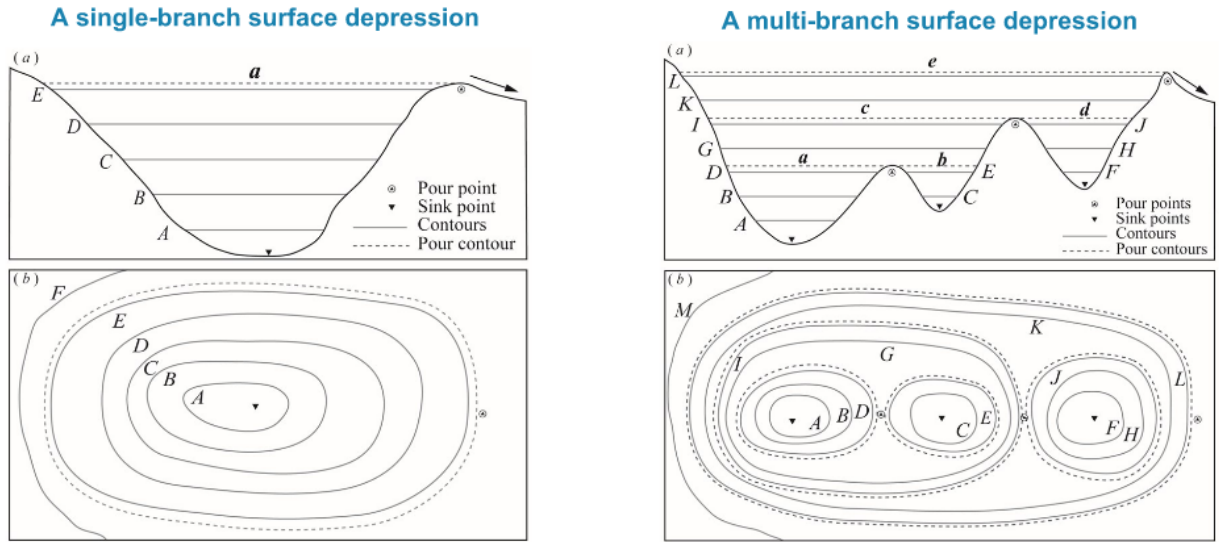
Note that the orientation field is an azimuth from 0-180 degrees, but that 0 is due east, and 90 is north-south, etc. If the sink is perfectly circular, the orientation value defaults to a value of 0. (Truthfully, the calculated orientation field may not be very useful given the various odd shapes of the depressions).

One can also determine the depth, area, and certain morphometric thresholds such as the index of circularity or eccentricity for sinks to retain. This can be done using a “training” dataset, or a polygon feature class that has been field-verified for actual depressions (e.g., Doctor and Young, 2013). At this point, distinguishing between actual sinkholes and other types of closed depressions is not necessary, as this is simply used to define thresholds for depressions in the elevation model.

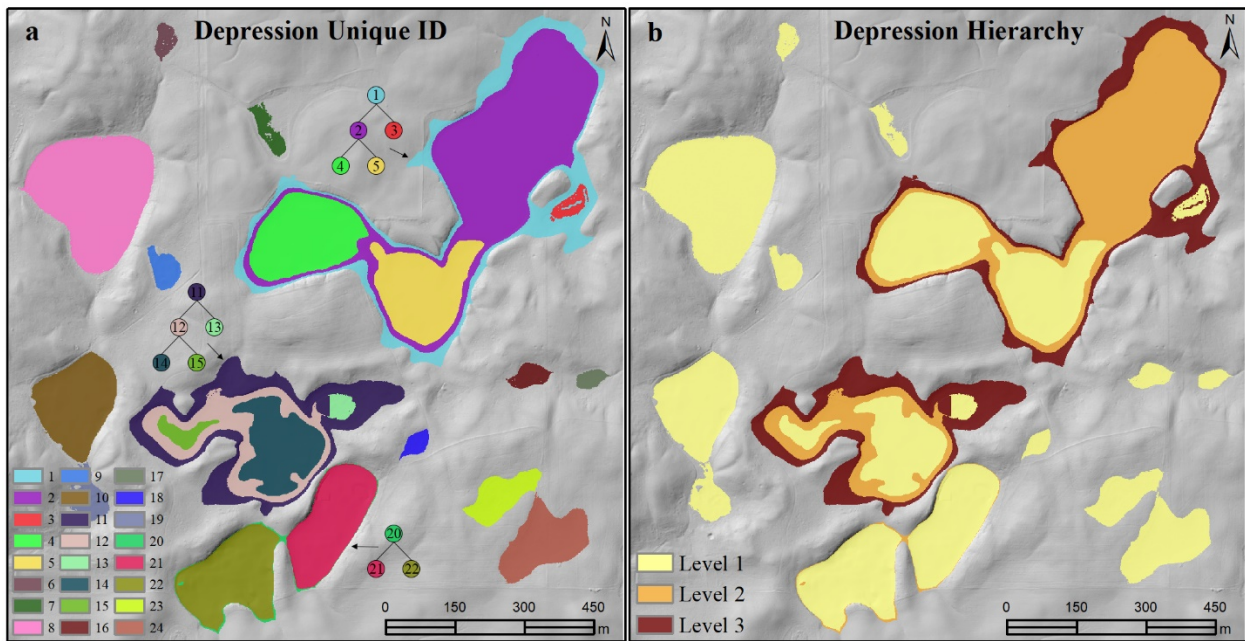
PART 2: Use the Contour Tree Tools for nested depressions

Tools for creating a nested hierarchy of closed contours and enclosing polygons were developed by Dr. Quisheng Wu at Binghamton University. In the following exercise, we will use Dr. Wu’s publicly available *Contour Tree Tools* toolbox for identification of nested depressions. For more information about him and his projects, please visit this website: <http://wetlands.io/#research>

There are two tools in the toolset: *Extract Sinks* and *Identify Depression Hierarchy*. The *Extract Sinks* tool selects areas of the DEM for contouring determined by a filling algorithm (much like the *Hydro* tool used above), so use of the *Contour Tree Tools* is best done on a burned DEM that has the “digital dams” at road/stream intersections already removed. The diagram below illustrates how the *Identify Depression Hierarchy* tool does the contouring that can identify progressive levels of nested depressions within a single larger one (Wu et al., 2015; 2016).

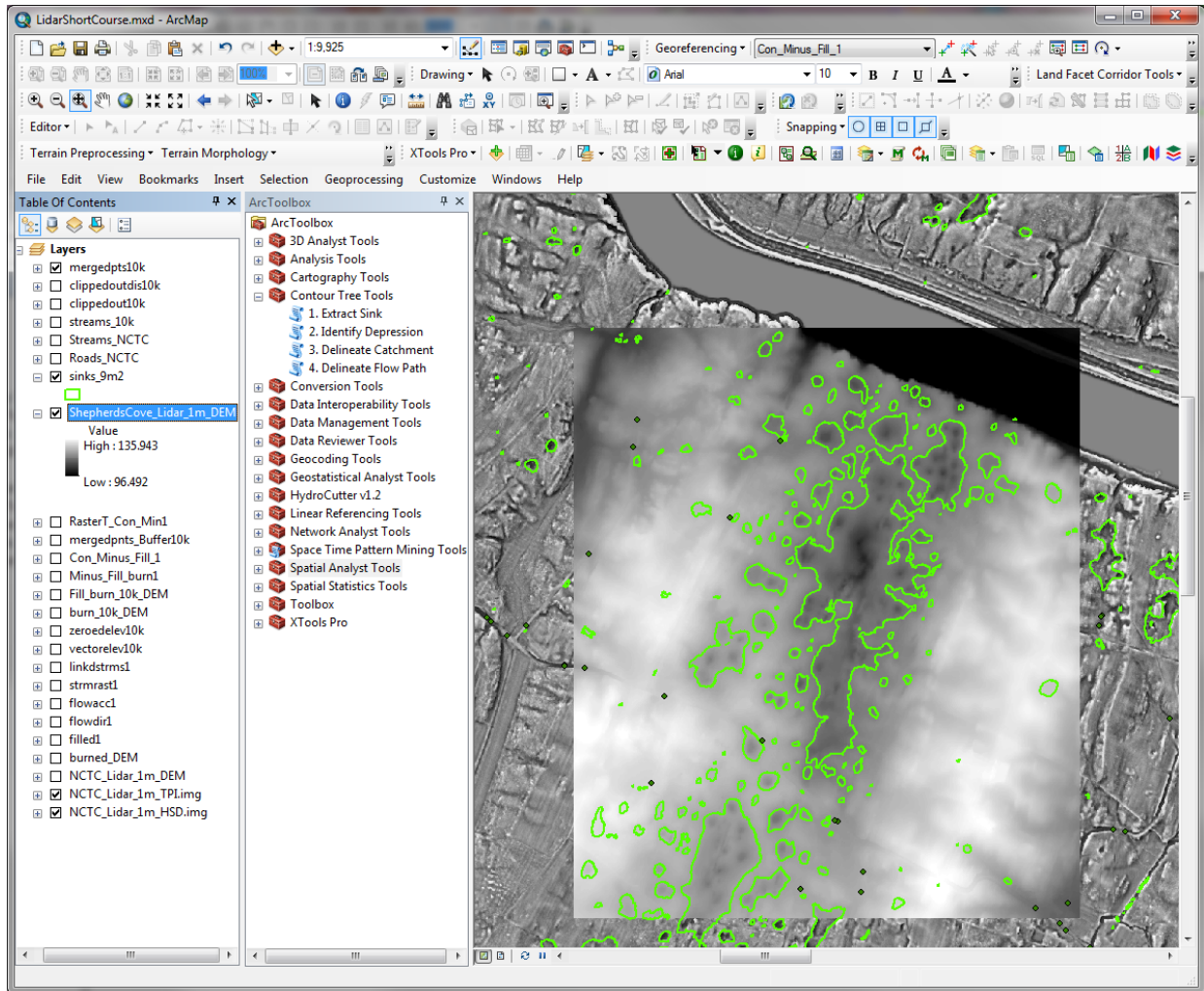


The *Identify Depression Hierarchy* tool also assigns a unique ID and non-unique “level” to each depression contour based upon how many additional nested depressions one might contain. Level 1 intervals contain no nested depressions, Level 2 contain more than one Level 1 depressions, Level 3 contain at least one Level 2 depression, and so on.

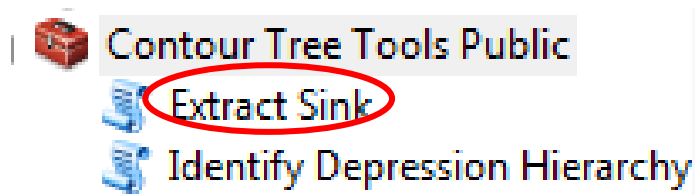


For this exercise, we will work on the area of concentrated sinkholes located within a region zoned for housing development called Shepherd's Cove that is located west of the NCTC center, and immediately south of the Potomac River. This site will be visited on the Friday field trip.

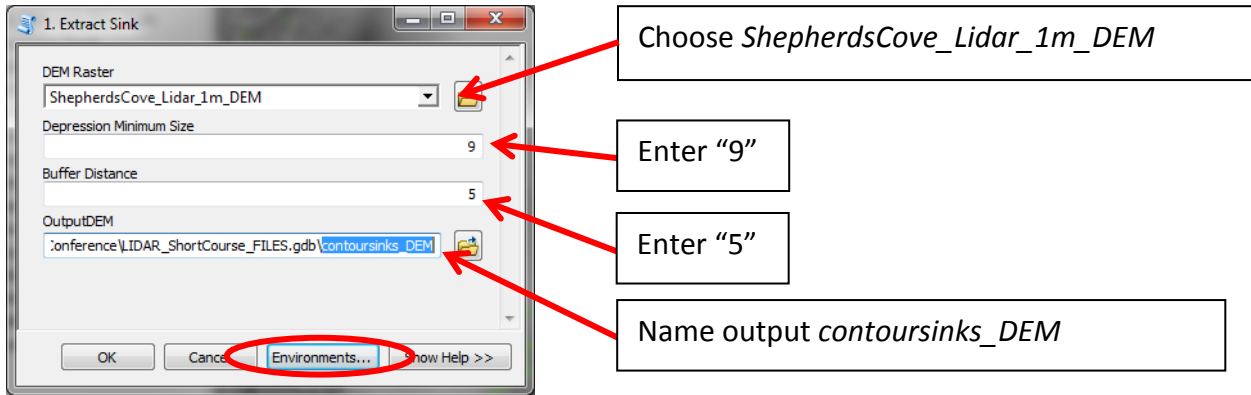
Add the *ShepherdsCove_Lidar_1m_DEM* to the project from the geodatabase. Right-click on the image and click **Zoom to Layer** (we will use this smaller area to speed up the processing time):



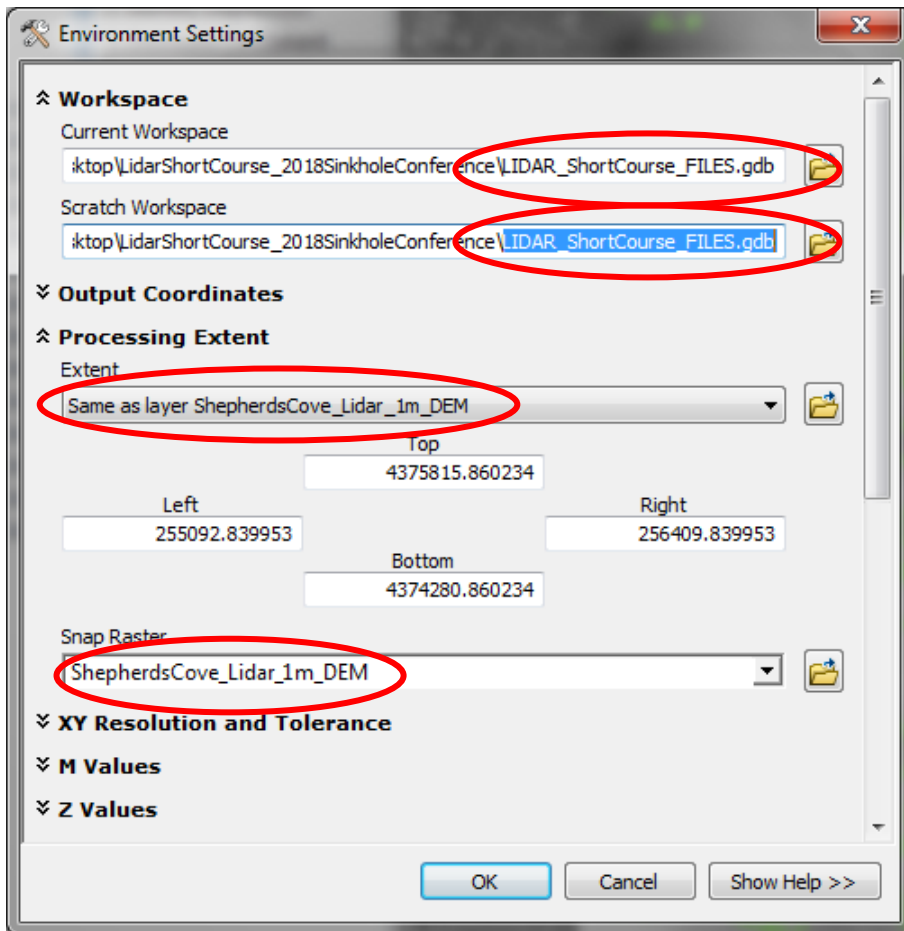
In the *Contour Tree Tools Public* toolbox, first run the *Extract Sink* tool to generate a subset of the DEM:



Set the **Depression Minimum Size** to “9” (the same area threshold used in the *Sinker* tool above). Set the **Buffer Distance** to “5”. Put the OutputDEM in the *LIDAR_ShortCourse_FILES.gdb* and name it *contoursinks_DEM*.

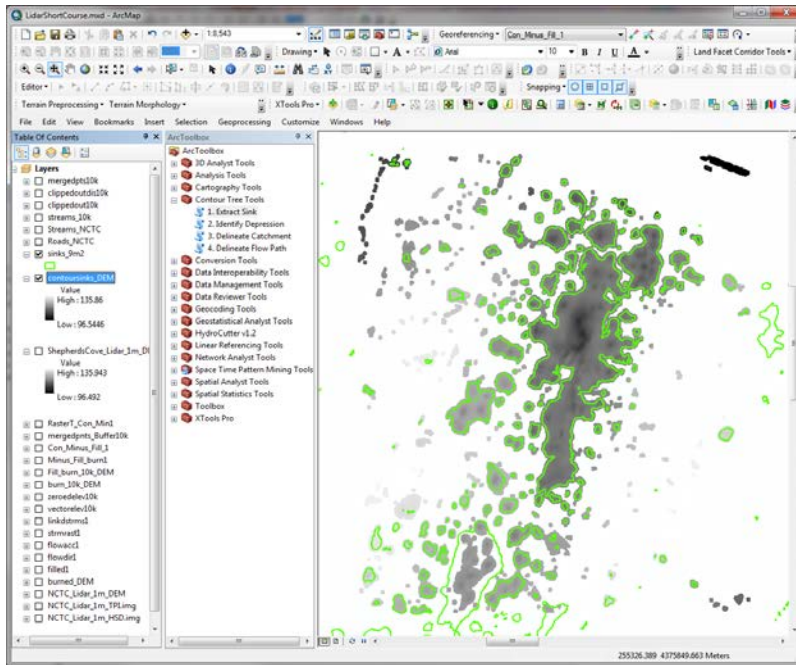


Click on the **Environments...** button, and update the **Current Workspace** and **Scratch Workspace** to the *LIDAR_ShortCourse_FILES.gdb*, and set the **Processing Extent** and **Snap Raster** to the *ShepherdsCove_Lidar_1m_DEM*:

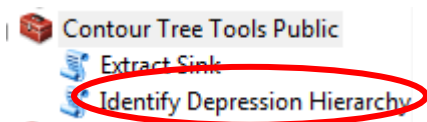


The tool should complete in less than 5 minutes. Once completed, turn off all layers except *sinks_9m2* polygons that had been generated earlier, and the new *contoursinks_DEM* layer.

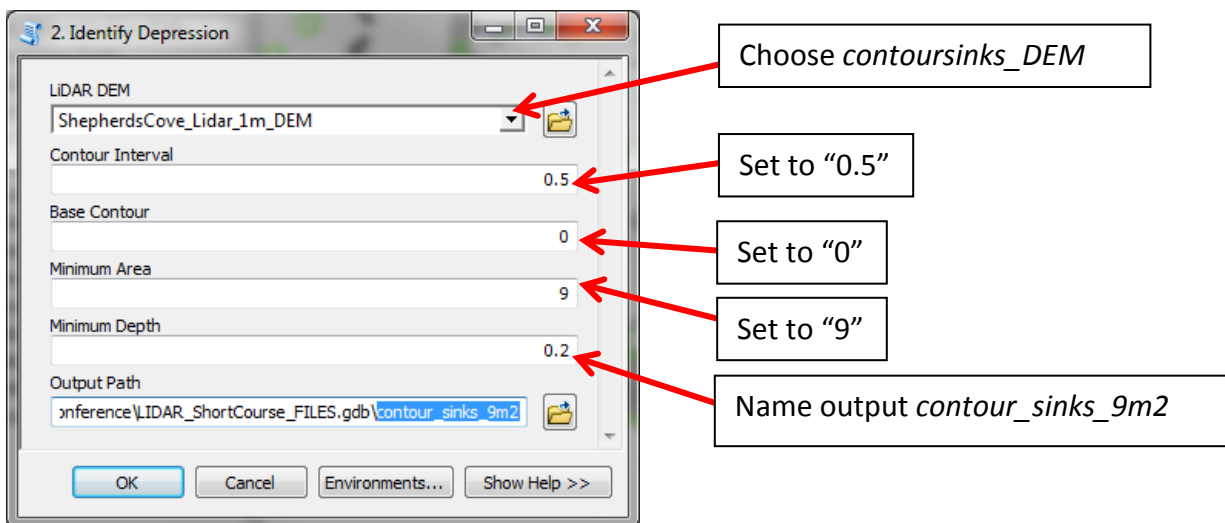
You should see something similar to the image below:



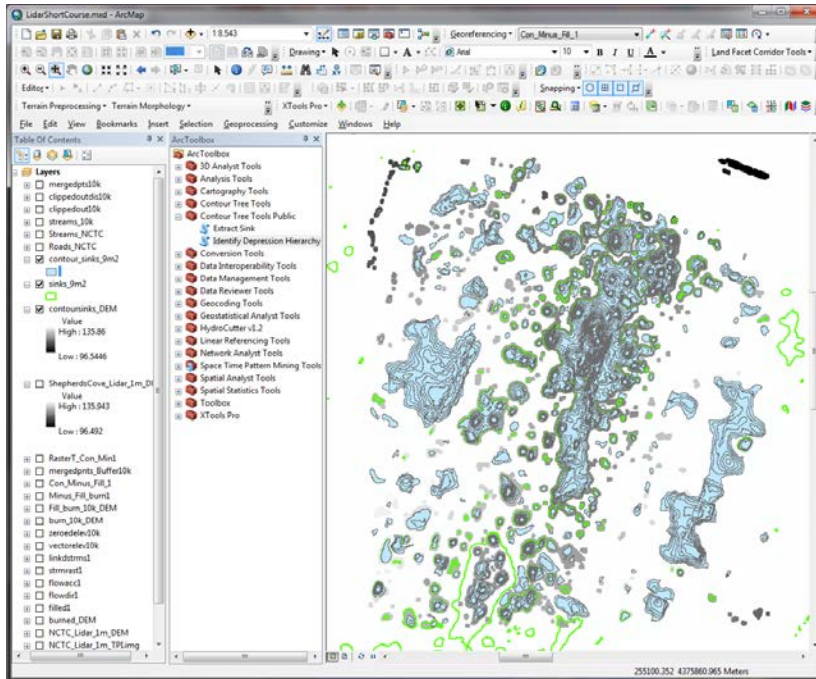
Next, run the *Identify Depression Hierarchy* tool



Set **Contour Interval** to “0.5”, **Base Contour** to “0”, **Minimum Area** to “9” (this is the same as the area threshold set earlier for the Fill-Difference method) and **Minimum Depth** to “0.2” (this is the same as the depth threshold set for the Fill-Difference method). The output path should be to *LIDAR_ShortCourse_FILES.gdb*, and name the output file *contour_sinks_9m2*.

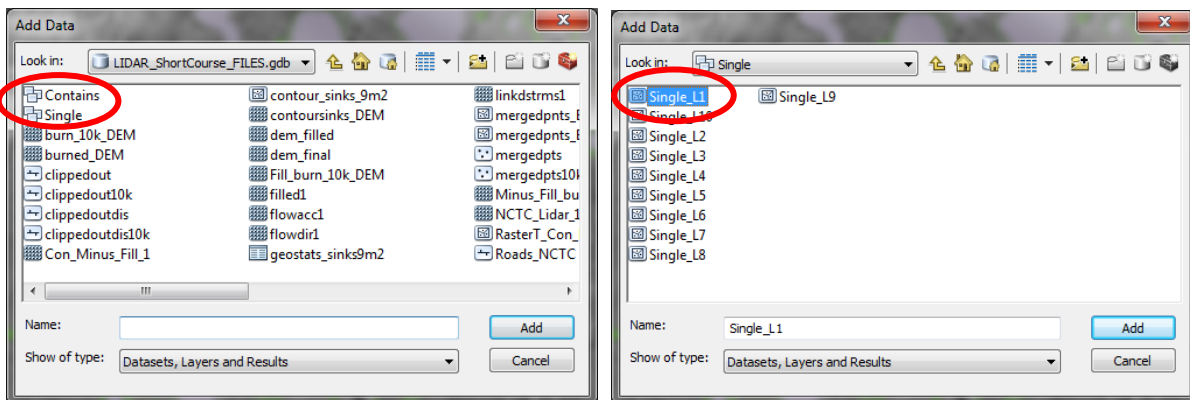


Click **OK** to run the tool. It should run for about 20 to 30 minutes. When finished, the results will look similar to the image below:



Add in the *NCTC_Lidar_1m_TPI.img* raster, and symbolize the *contour_sinks_9m2* polygons to be without a fill color and with red outlines.

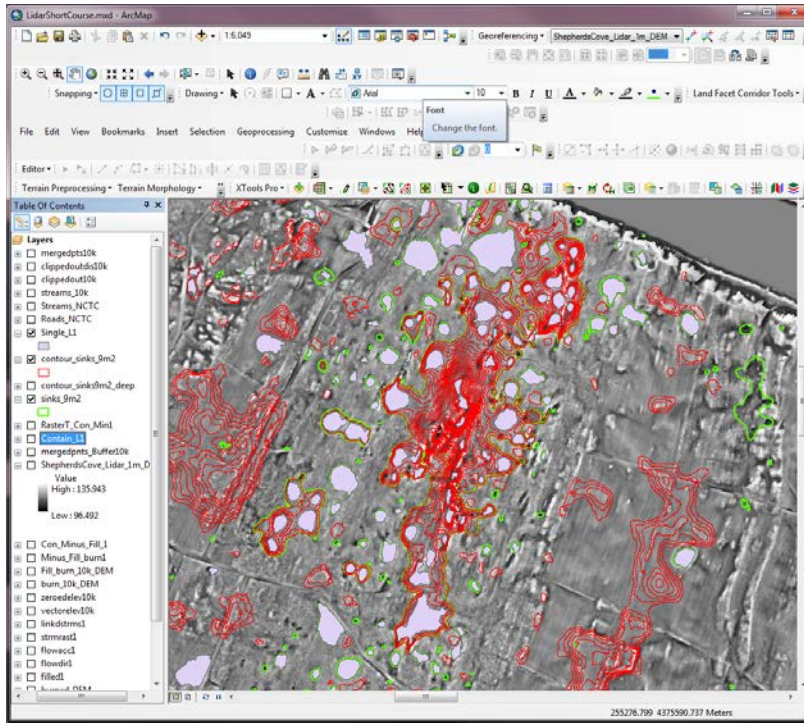
Next, go to **Add Data** → and within the geodatabase you should see two new feature datasets, *Contains* and *Single*. Open the *Single* feature dataset, and add the feature class **Single_L1**:



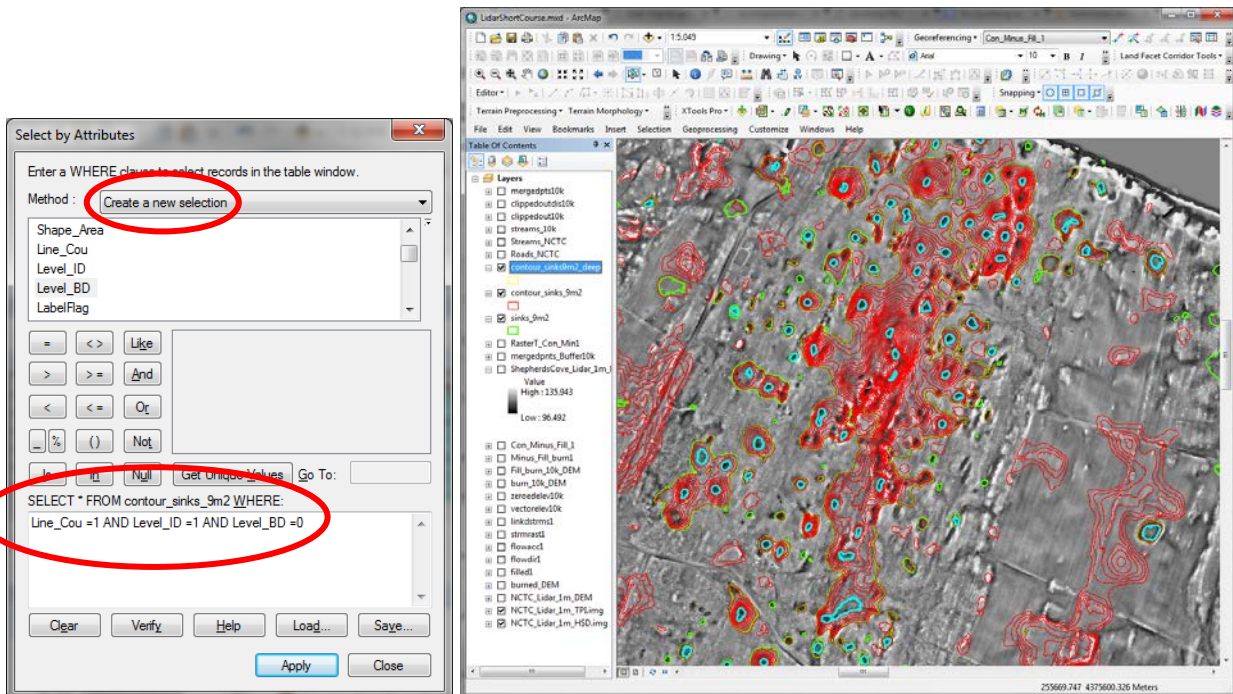
Compare the results of the *Contour Tree Tool* to the results obtained earlier with the Fill-Difference method. Note that the *Single_L1* depressions are those nested within larger depressions, but do not contain other nested depression levels themselves. This dataset can be useful for defining an overall depression inventory in areas of compound depressions.

Explore the other levels of the *Single* feature dataset. Open the attribute table(s) of each. Note that the following attributes were written to each depression: the outermost contour elevation (*Cont_ field*), the

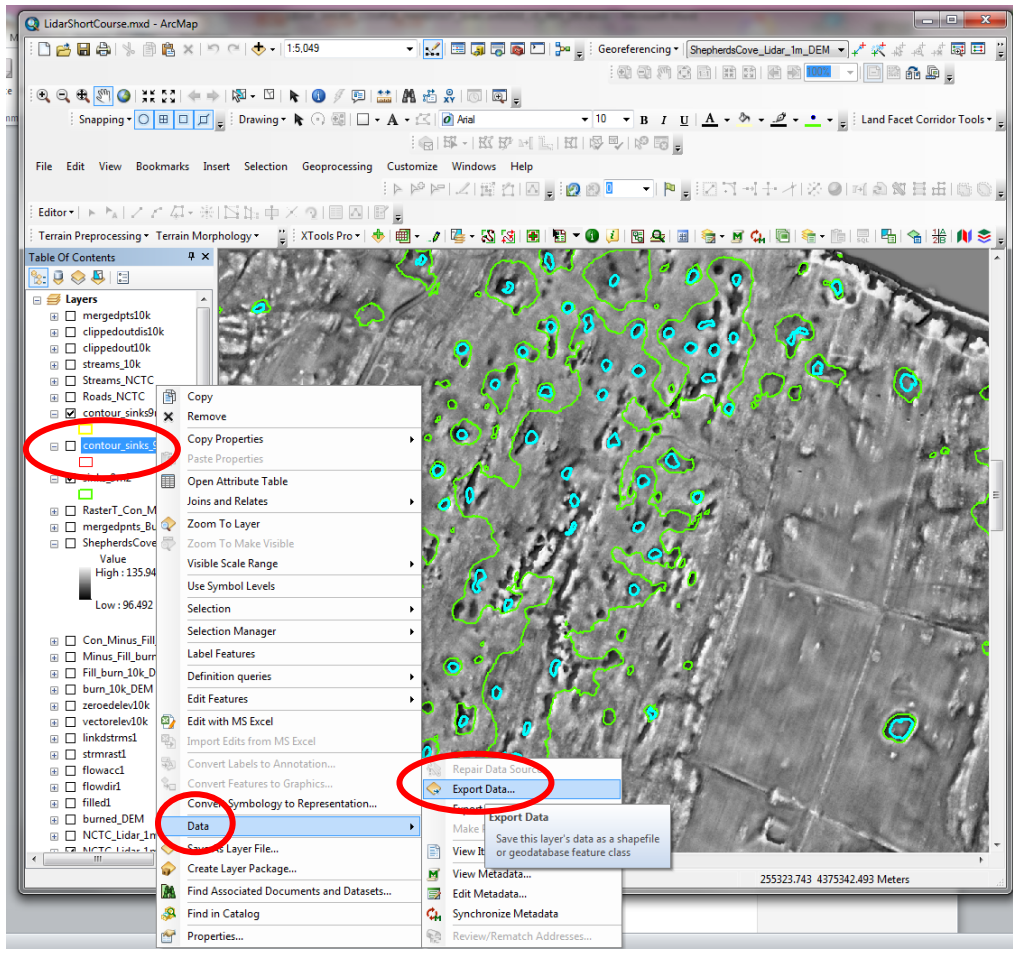
area, the min, max, range, mean and standard deviation of elevations, the sum of all elevation values, and the volume of the depression in cubic meters.



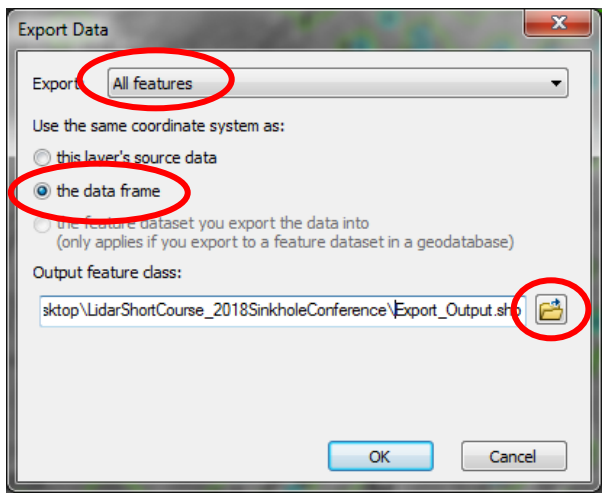
You can also identify the lowermost Level 1 contours within the compound sinkholes of the *contour_sinks_9m2* polygons (i.e. the lowest areas in the bottom of multiple-contoured depressions) by selecting according to the attributes of the table where “Lin_Cou = 1 AND Level_ID=1 AND Level_BD =0”:



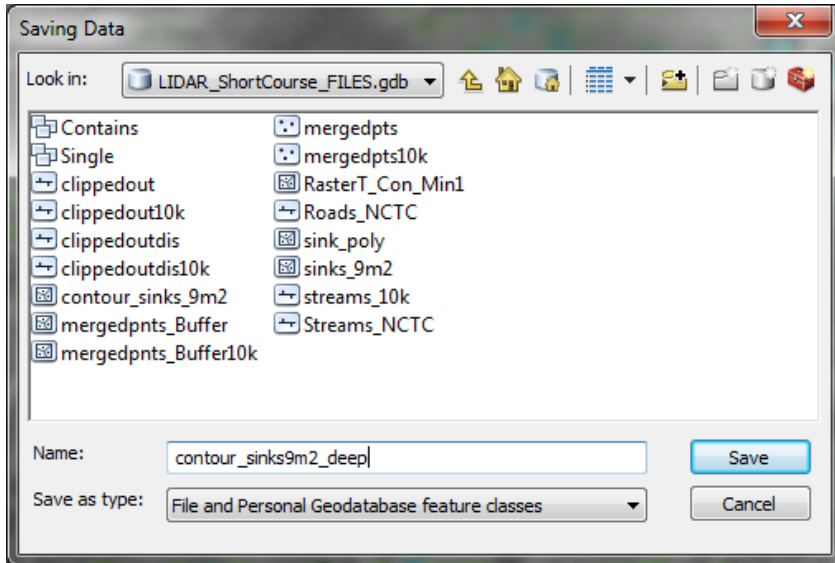
The resulting selected output could be useful as a layer of the deepest local sinks in a region of nested sinks. To create a standalone dataset from the selected deepest depressions nested within larger ones, right click on the layer `contour_sinks9m2_deep`, then select **Data** → **Export Data**.



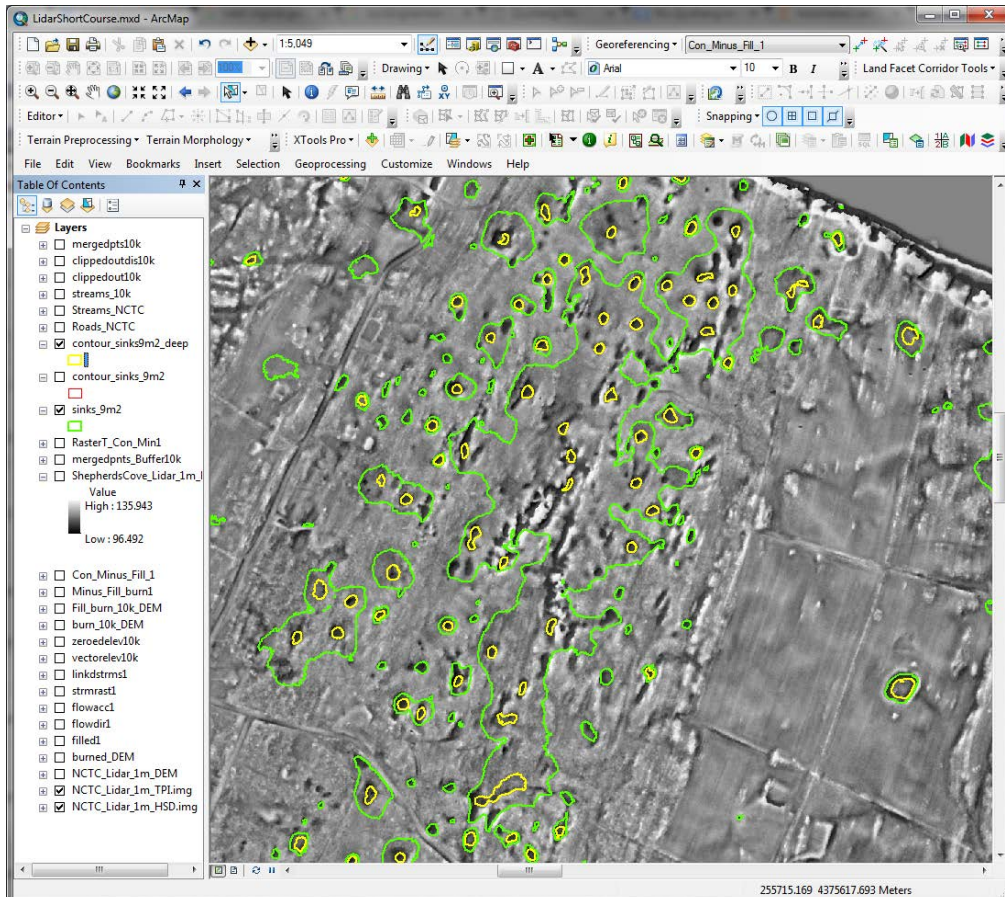
Export *All features*, and select *the data frame* to use as the coordinate system of the dataset. Next, click on the folder for the output location...



Under *Save as type*, select *File and Personal Geodatabase feature classes*, and the geodatabase icon should appear. Open the geodatabase, and rename the file *contour_sinks9m2_deep*.



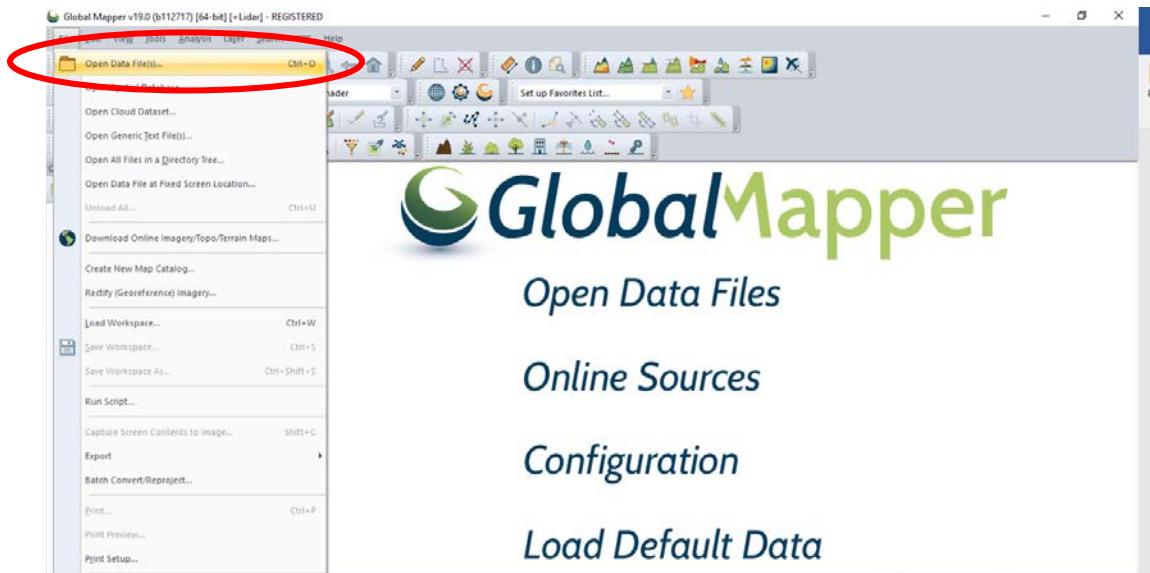
Allow the data to be added to the current map, and re-symbolize the outlines of the new polygon layer to be yellow. Turn off all depression layers except *sinks_9m2* and *contour_sinks9m2_deep*. You should see something like the image below:



EXPORT TO GOOGLE EARTH:

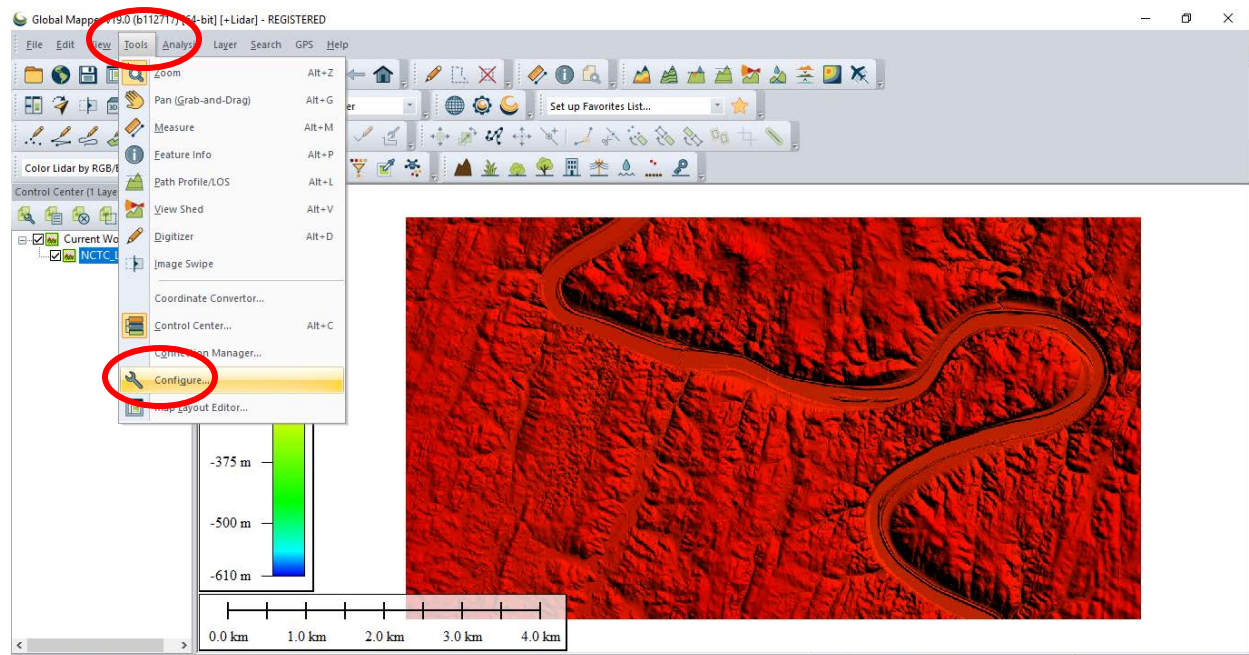
It is often useful to overlay the lidar imagery in Google Earth to explore, or to share with others. ArcGIS is not the optimal software for exporting large amounts of imagery to Google Earth; for this purpose, we recommend using Global Mapper.

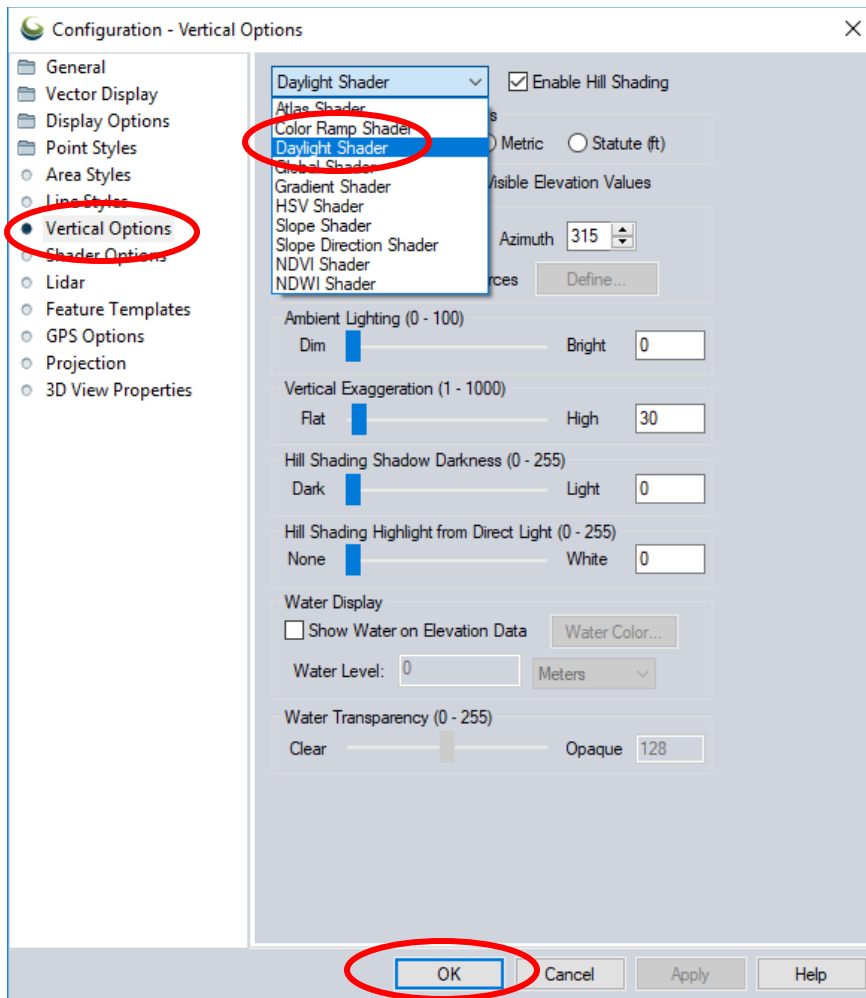
Open the Global Mapper software:



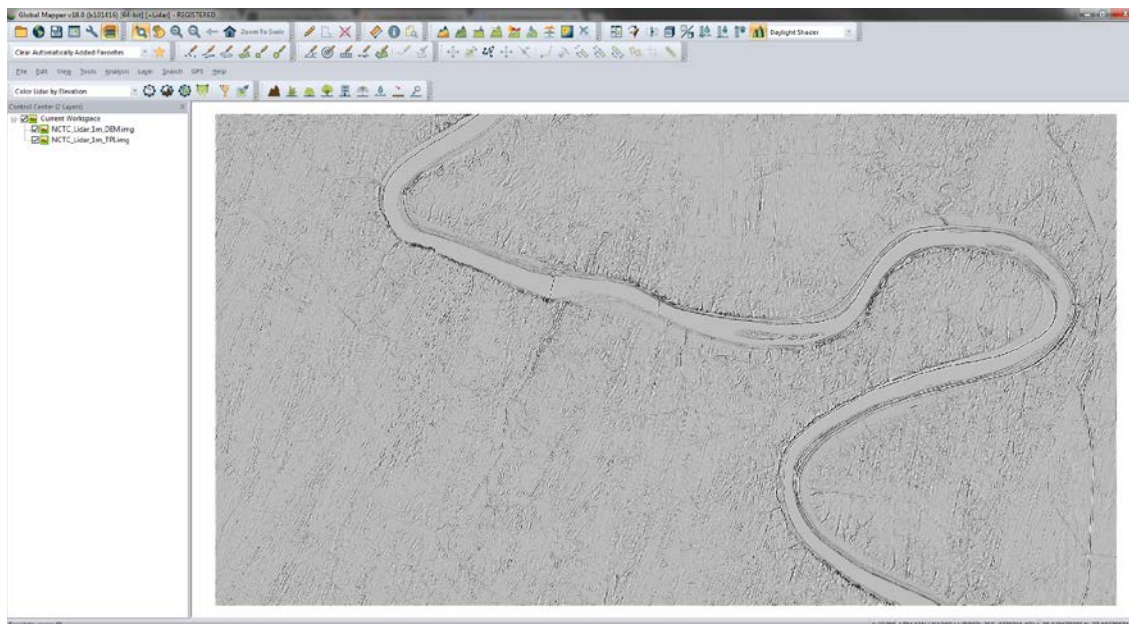
Add the *NCTC_Lidar_1m_DEM.img* file from the folder on the desktop. This DEM is NOT located within an ESRI geodatabase.

You should see something similar to what is shown below. Click on **Tools** → **Configure**, click on **Vertical Options**, and select the **Daylight Shader**.

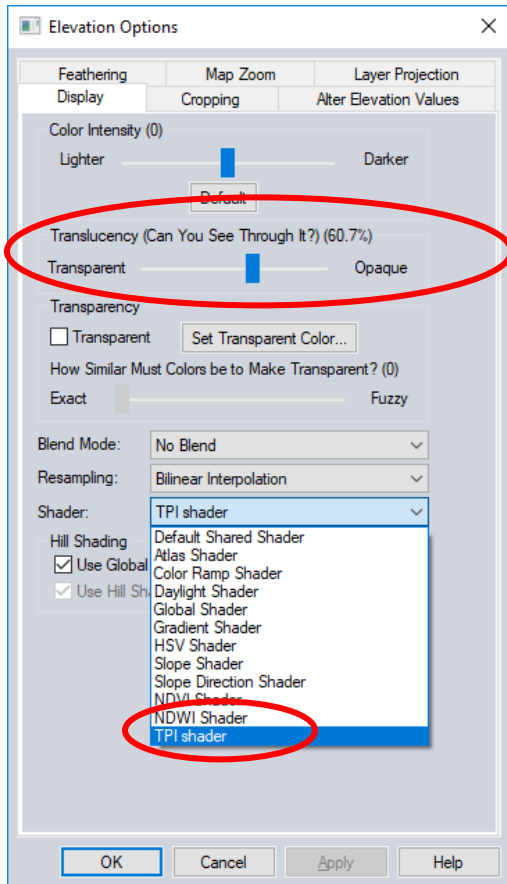




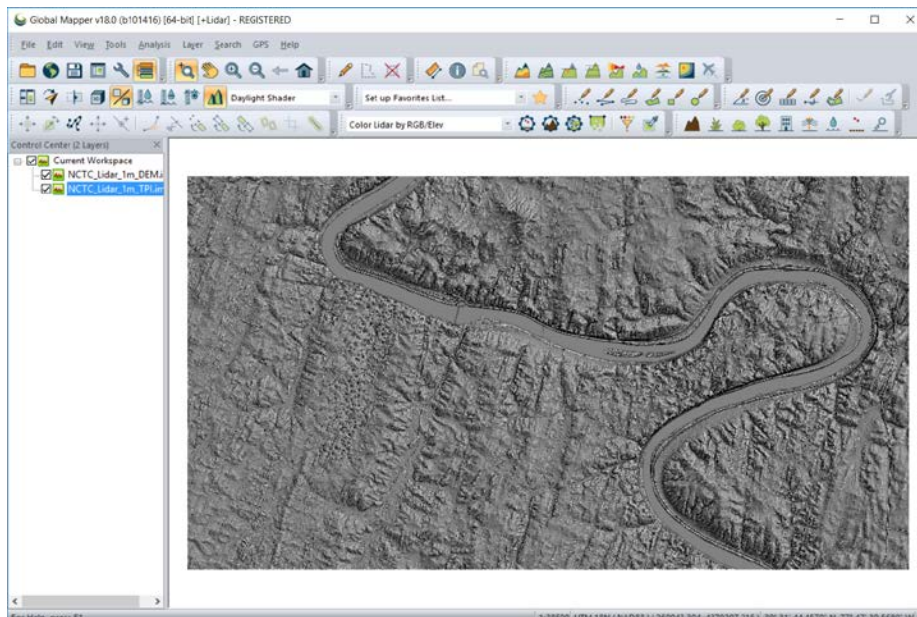
Add the *NCTC_Lidar_1m_TPI.img* image you created earlier, to see something like this:



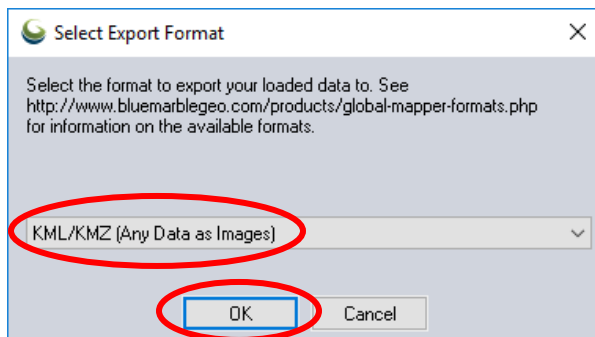
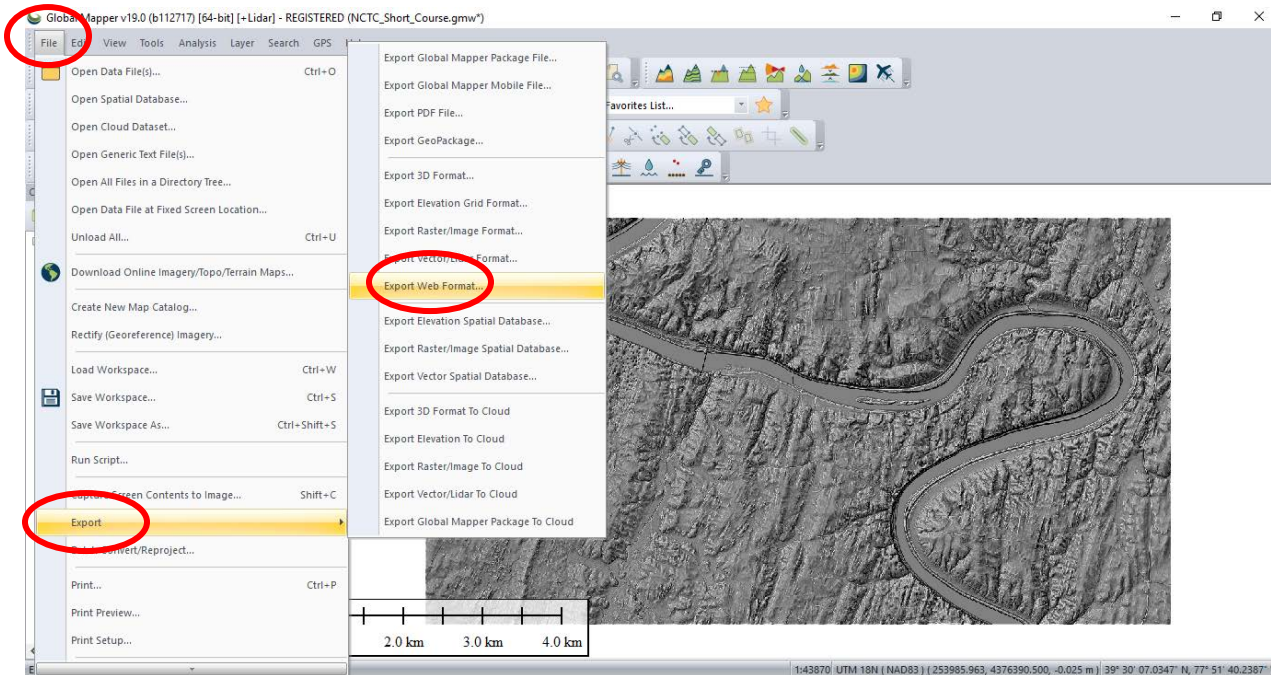
Right-click on the file name *NCTC_Lidar_1m_TPI.img*, and select **Options**. Under the **Display** tab, change the **Translucency** slider to ~60.0%, and under drop down list for **Shader** select *TPI Shader*. This is a customized shader option for displaying the TPI image appropriately.



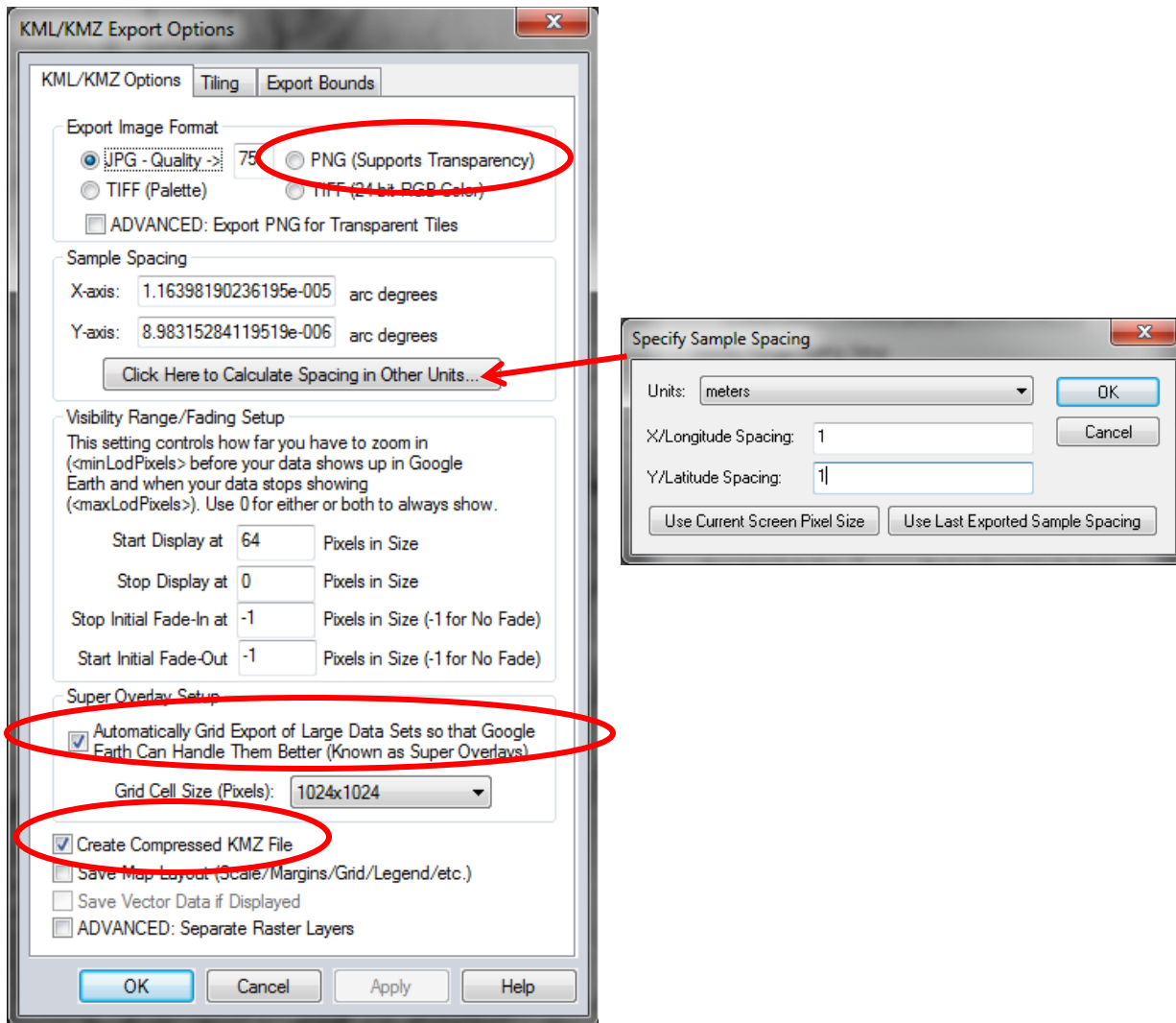
The resulting image should look something like that below:



Next, go to **File** → **Export** → **Export Web Format...** select the **Export Format** to be *KML/KMZ (Any data as images)*.



Under the KML/KMZ Options tab, select an **Export Image Format** of *PNG (Supports Transparency)*. Check that the **Sample Spacing** and note that it is in arc degrees. For our visualization, we want 1 m so click on **Click Here to Calculate Spacing in Other Units...** For **Units** select *meters* and for **X/Longitude Spacing** and **Y/Latitude Spacing** enter “1”. Click **OK**. Check the box under **Super Overlay Setup**, and leave the default **Grid Cell Size (Pixels)** at *1024x 1024*. Under **Altitude Mode** (if shown, depending on the version of Global Mapper being used), select *Clamp to Ground*—this is the default setting for newer versions of Global Mapper. Click **OK**, and save the file as *NCTC_Lidar_1m_TPI+HSD* to the desktop folder.



After some time to process the multi-layered tiled imagery, a KMZ file will be saved that can be opened in Google Earth.

REFERENCES:

Doctor, D.H. and Young, J.A. (2013). An evaluation of automated GIS tools for delineating karst sinkholes and closed depressions from 1-meter LIDAR-derived digital elevation data. In: Land L, Doctor DH, Stephenson JB, editors, Sinkholes and the Engineering and Environmental Impacts of Karst: Proceedings of the Thirteenth Multidisciplinary Conference, May 6-10, Carlsbad, New Mexico: NCKRI Symposium 2. Carlsbad (NM): National Cave and Karst Research Institute, p. 449-458.

Wall, J., D. H. Doctor, and S. Terziotti (2015) A Semi-automated Tool for Reducing the Creation of False Closed Depressions from a Filled LIDAR-derived Digital Elevation Model. In: Doctor. D.H., Land. L., and Stephenson, J. Brad, eds., National Cave and Karst Research Institute Symposium 5, Proceedings of the 14th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst, Oct. 5-9, Rochester, Minnesota, pp. 255-262.

Wu, Q., Deng, C. and Chen, Z., 2016. Automated delineation of karst sinkholes from LiDAR-derived digital elevation models. *Geomorphology*, 266, pp.1-10.

Wu, Q., Liu, H., Wang, S., Yu, B., Beck, R. and Hinkel, K., 2015. A localized contour tree method for deriving geometric and topological properties of complex surface depressions based on high-resolution topographical data. *International Journal of Geographical Information Science*, 29(12), pp.2041-2060.

APPENDIX 1: Obtaining lidar and other geospatial data served by the USGS

The “bare earth” lidar DEM and the datasets of stream line vectors and road vectors have already been provided for you for use in these short course exercises in a ESRI geodatabase within the folder ‘LidarShortCourse_2018SinkholeConference’ on your desktop. The following instructions are for future use as a guide for obtaining publicly available lidar data.

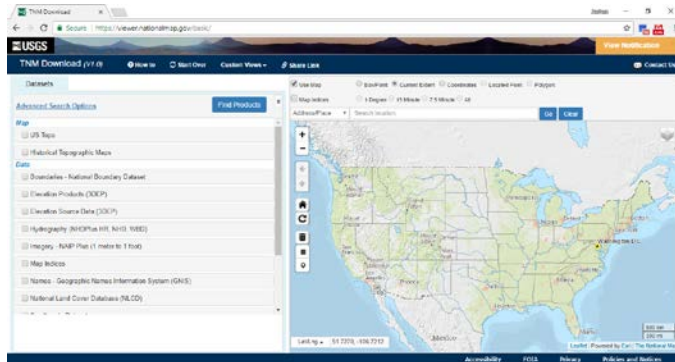
A continually-updated inventory of publicly available lidar elevation data is served through the United States Interagency Elevation Inventory, and can be viewed online at this website:

<https://coast.noaa.gov/inventory/>

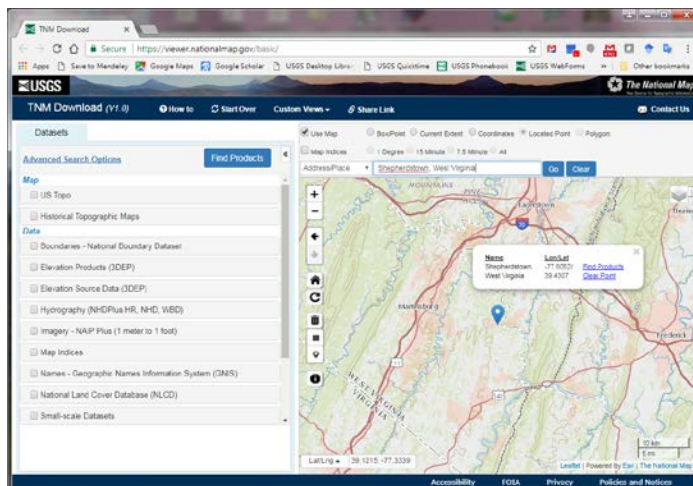
The USGS serves digital elevation data for the nation produced by the 3D Elevation Program (3DEP) through the web portal of the National Map. To get the data, go to:

<https://viewer.nationalmap.gov/basic/>

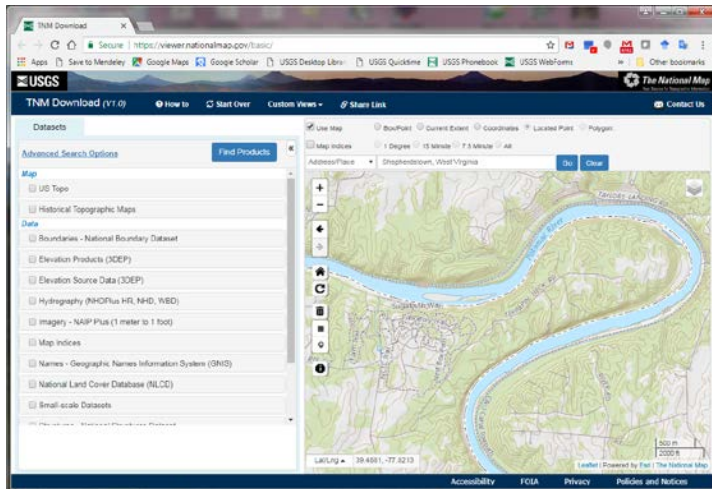
This will launch a web-based viewer and download platform as shown below:



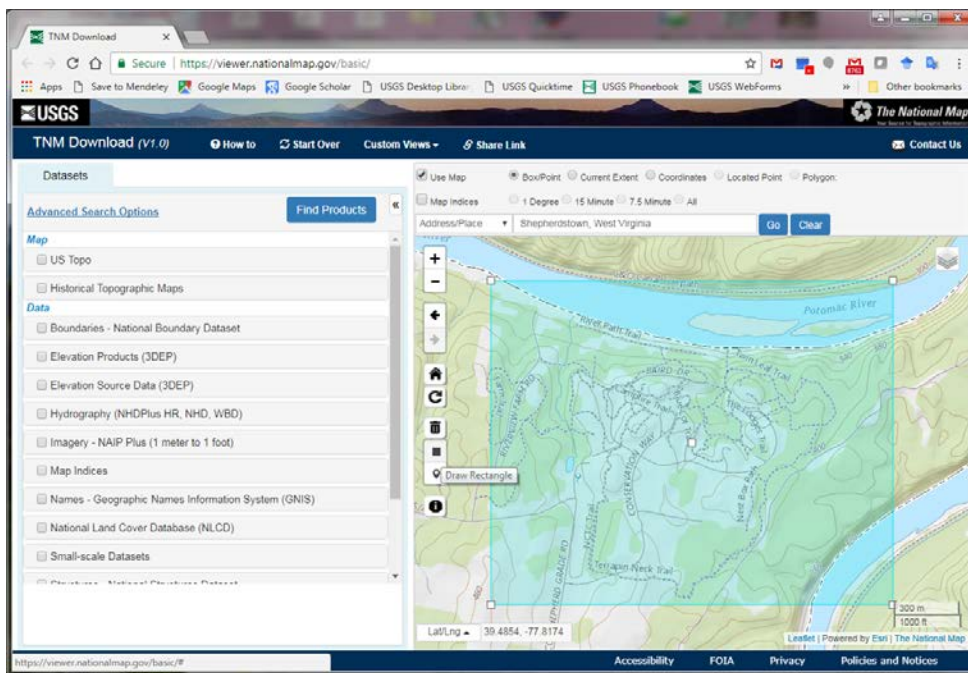
Type “Shepherdstown, WV” in the box that says Search Location:



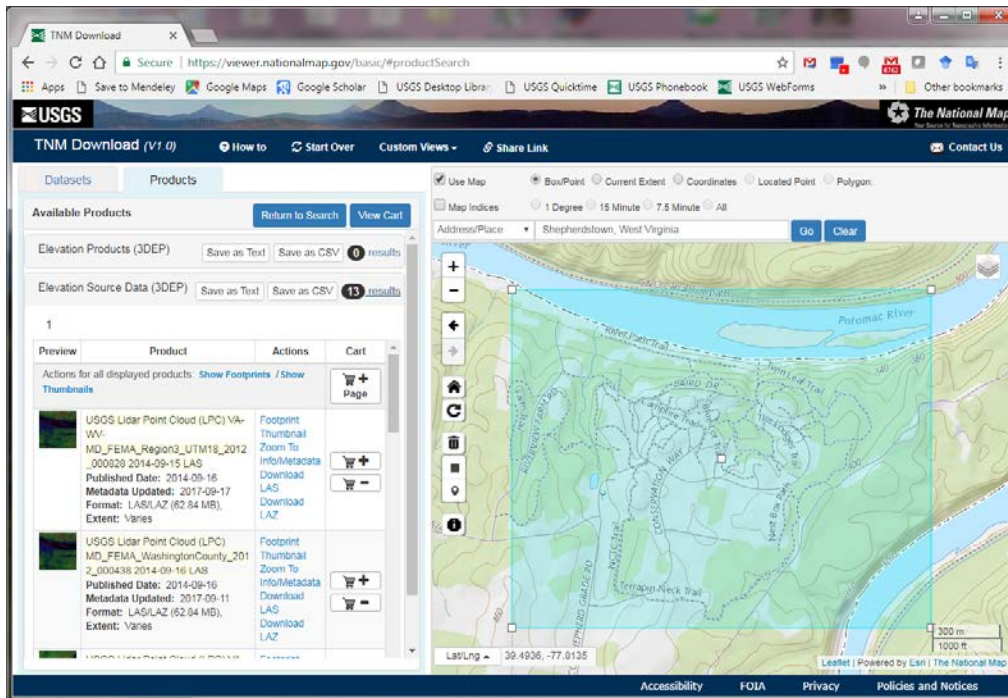
Zoom in to the meander bend of Terrapin Neck just below Mercerville, MD to focus on the site of the NCTC:



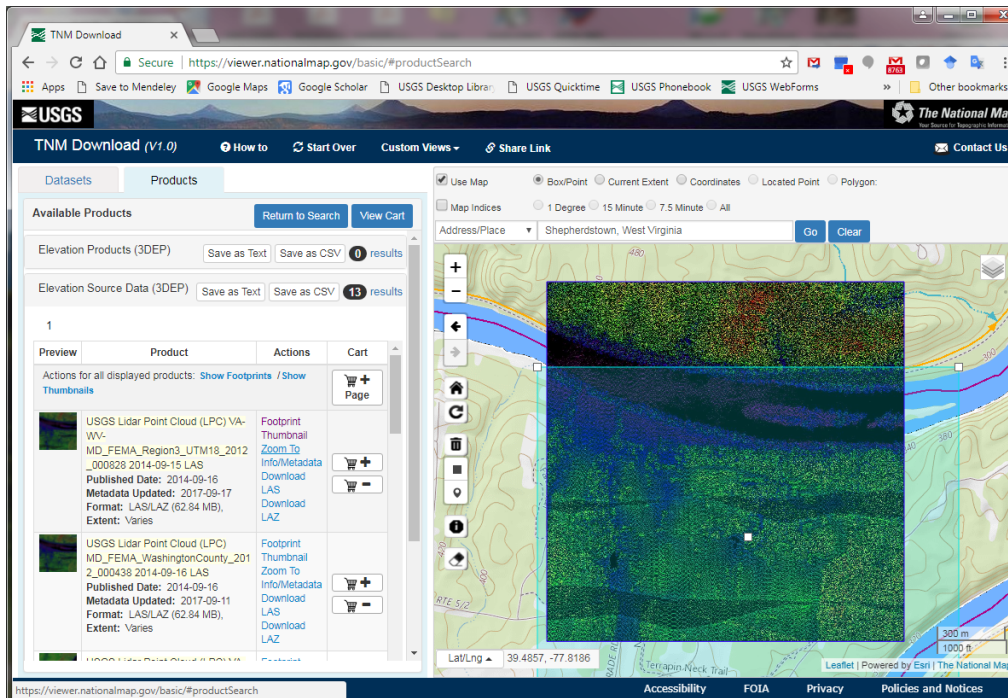
Click the square **Draw Rectangle** icon and draw a bounding box to highlight the area around the NCTC.



In the left-hand menu, check the box **Elevation Products (3DEP)**, uncheck the **1/3 arc-second DEM** option, and then check the box that says **1 meter DEM**. Scroll down and check the box **Elevation Source Data (3DEP)** and check the box **Lidar Point Cloud (LPC)**, then click the blue button at the top **Find Products**. A list of all available data products in the region of interest will be shown.

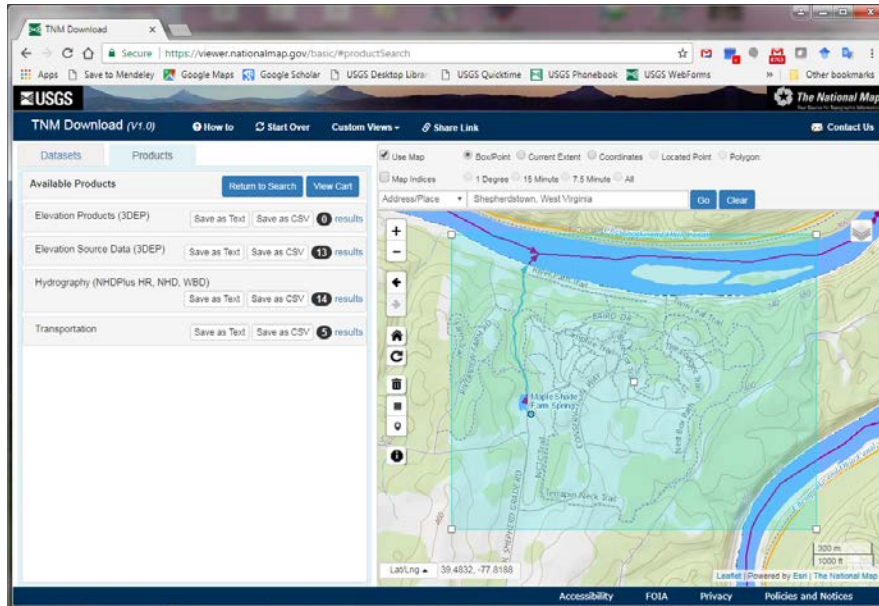


Note that for this region, the 1 meter resolution “bare earth” DEMs are not served; however, the original lidar point cloud data are available as 1.5 km x 1.5 km tiles.



In order to create a DEM, software capable of reading LAS/LAZ file formats for creating an elevation surface from the point cloud data would be needed. Although processing of point cloud data to create a DEM is beyond the scope of this course, the DEM has already been processed and is provided for your use later in the course.

Return to the **Datasets** tab, scroll down to **Hydrography**, and select **National Hydrography Dataset (NHD)**. Then, scroll down to **Transportation** and select **National Transportation Dataset**. Click on **Find Products** again, and you will see the hydrography and transportation data added to the list.



To download these or other digital geospatial products in the future, simply add the products you wish to obtain to the cart, and follow the instructions for downloading them (no login required). Downloads of large amounts of data can be facilitated using the USGS **uGet** utility (<https://viewer.nationalmap.gov/uget-instructions/>).

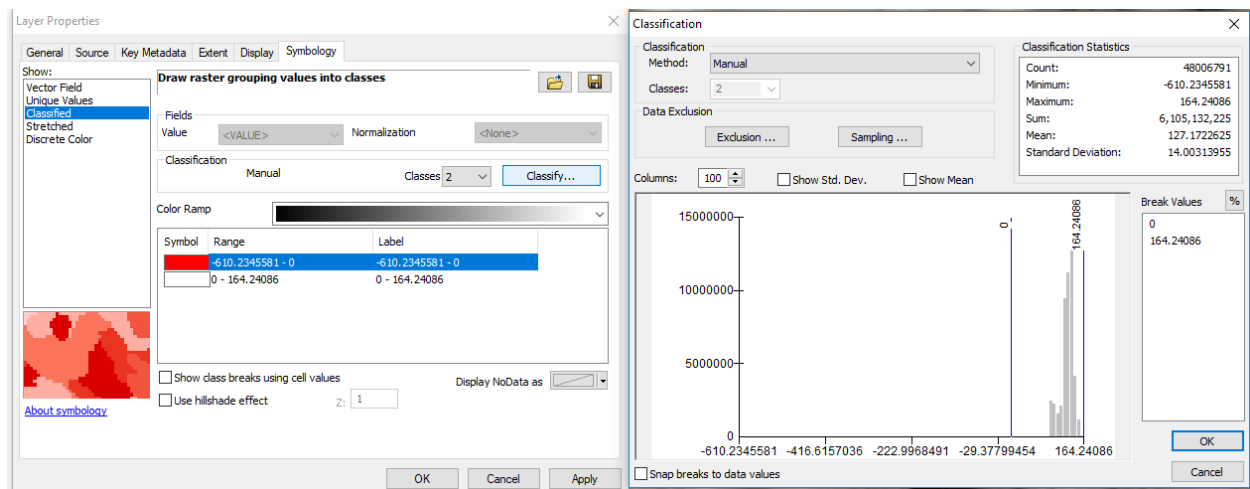
APPENDIX 2: How to quickly evaluate elevation data


Finding erroneous pixels

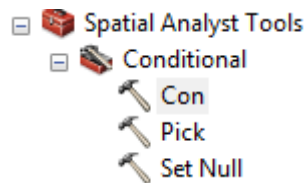
Here we quickly review how to evaluate elevation data. Looking at the elevation data that has been provided, note the values range between -610.235 and 164.241. Values of -610 seem erroneous; so, let's make sure we know where these values are before we do anything.



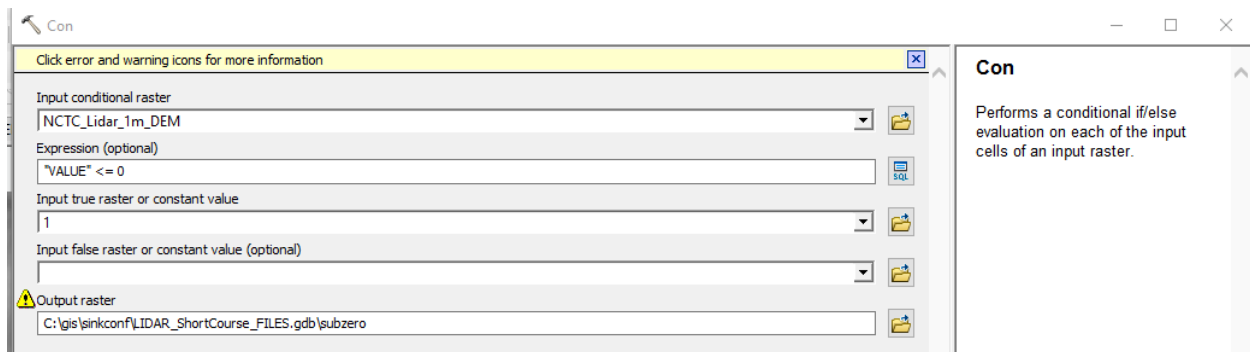
Double-clicking on the title for the DEM will pull up the **Layer Properties**. Click on the **Symbology** tab and then select **Show → Classified**. Set the **Classes** to "2" then click on the **Classify...** button. Set your break value for the first class to be zero. Note that the histogram of elevations has very few values below zero. Press **OK** to accept the **Break Values**. Double clicking on the **Symbol** color next to the first class, select something that will stand out like *Mars Red* (bright red) while setting the color for the second class to something easily avoided like **No Color**. Click **Apply** and note that your screen appears to turn white. In this instance, there are so few pixels below zero that your analysis is unlikely to be impacted.



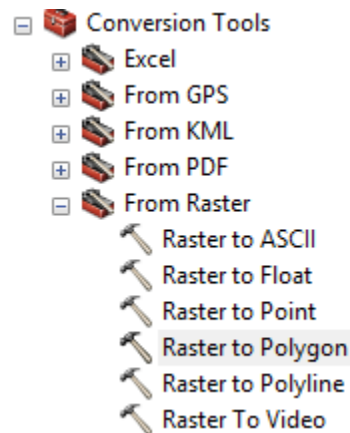
To go more in-depth, we can convert the values that are sub-zero to polygons to determine the amount of area that has these values. To do this, within the **ArcToolbox** () navigate to **Spatial Analyst Tools** → **Conditional** and select the *Con* geoprocessing tool



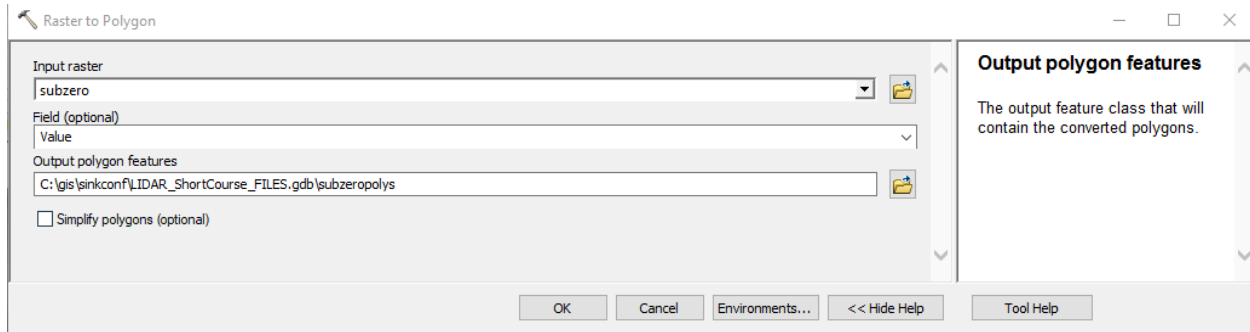
Set the input conditional raster to *NCTC_Lidar_1m_DEM*, set the **Expression** to “*VALUE*” <= 0, and then the input true raster or constant value of “1”. Set the output to something reasonable such as *subzero*.



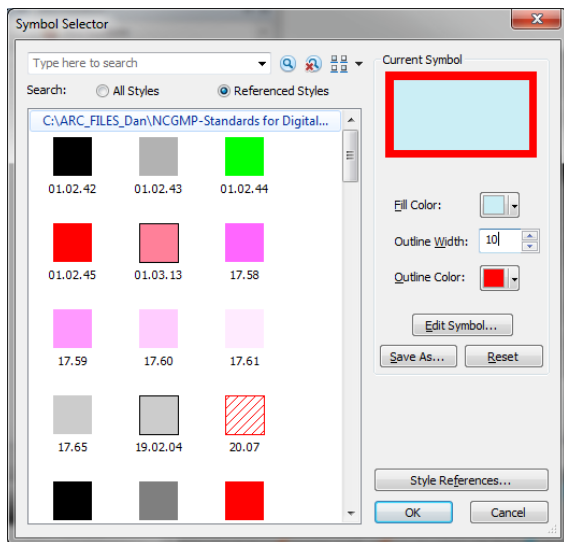
All pixels less than or equal to zero now have a value of “1.” Within the **ArcToolbox**, navigate to **Conversion Tools** → **From Raster** and select the *Raster to Polygon* geoprocessing tool



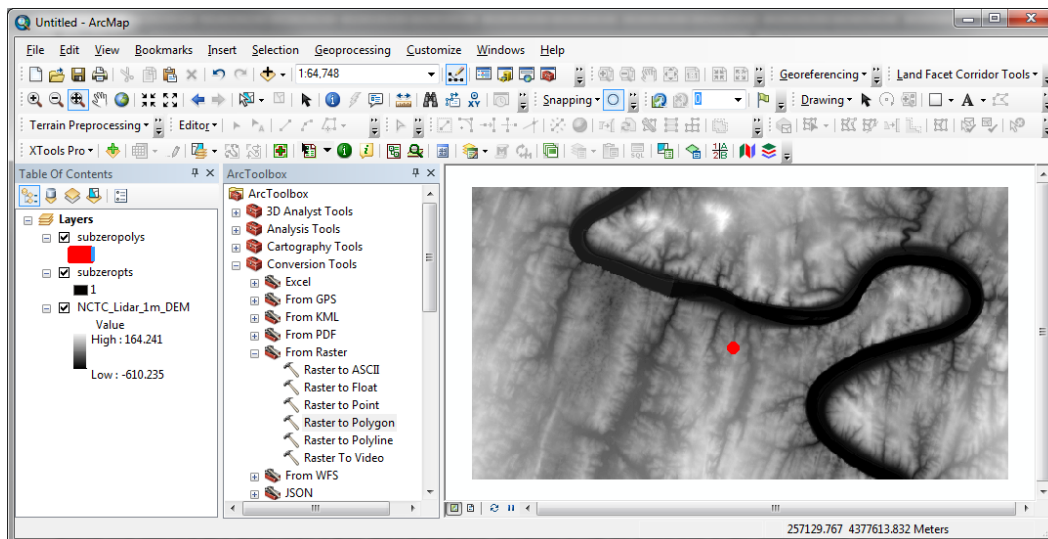
Set the Input raster to your previous output raster (i.e. *subzero*). Note that **Field** should auto-populate with **Value**. Set the output to something descriptive such as *subzeropolys* and then untick **Simplify polygons (optional)**.



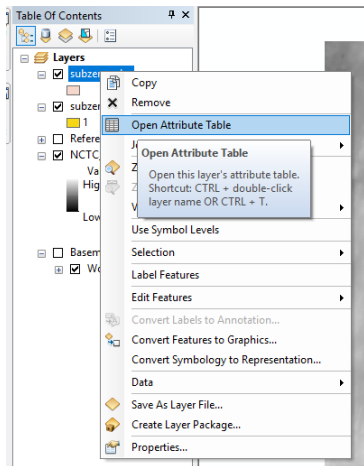
Now that a polygon layer with the pixels of sub-zero elevation has been created, we can change the symbology in order to see them. Click on the polygon symbol in the **Table of Contents**, then use the **Symbol Selector** tool to change the **Outline Color** of the polygon to a *Mars Red* and an **Outline Width** of 10:



Now, you should be able to see where the offending pixel sits within the DEM:



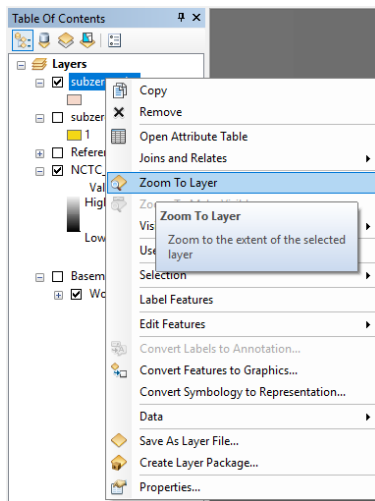
Since we placed the data into a File Geodatabase, metrics such as area, are auto-calculated. Right-click on the polygon layer and select **Open Attribute Table**.



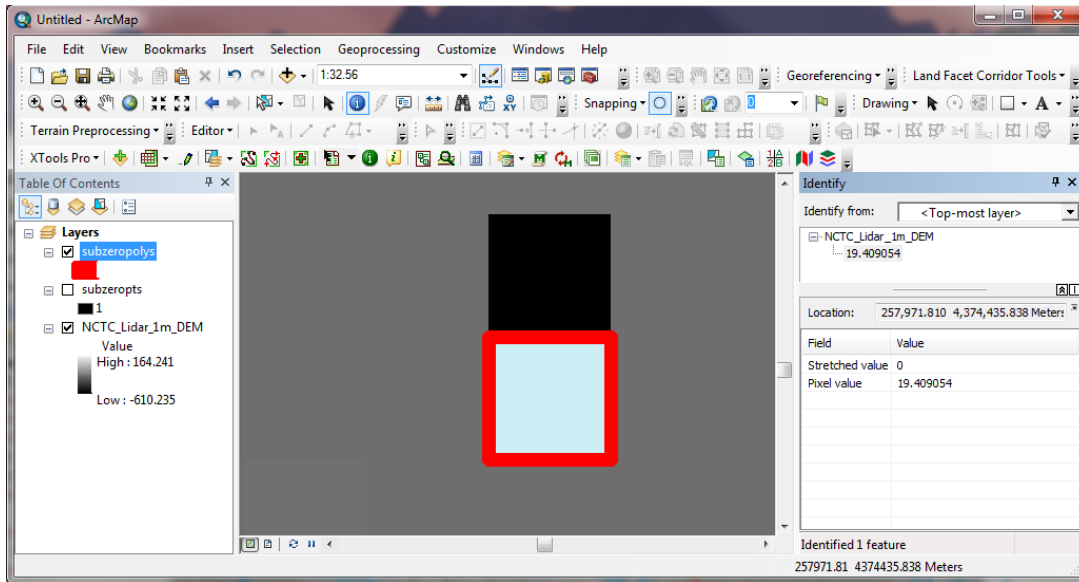
Note the *Shape_Area* column has a value of “1” meaning that there was only one pixel in the DEM that had a value of less than or equal to zero.

OBJECTID	Shape	Id	gridcode	Shape_Length	Shape_Area
1	Polygon	1	1	4	1

Right clicking on *subzeropolys*, select **Zoom To Layer**




You will then be taken to the offending pixel. Note that there is a second pixel (shown black in the figure below) located next to the subzero pixel that has an anomalously elevation value (+19.409054 meters) where surrounding values are over 100 meters above it. This might have been a result of the methods use to generate the DEM or it could represent a location where there were too few lidar points to adequately interpolate the data.




To fix the erroneous pixels in the DEM, you can reclassify any pixels with negative elevations to be “Null” values using the *Set Null* tool in **Spatial Analyst Tools** → **Conditional**.

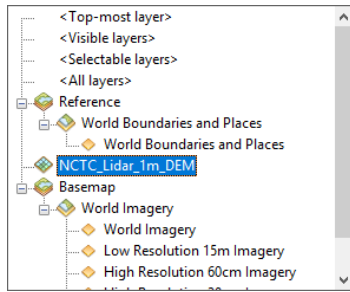
Validating vertical units

You will note that when the elevation ranges were listed above there were no units. While we might assume they are in our favorite units, it is always a good idea to verify this.

Change the layer properties back to a stretched black-to-white color ramp. Clicking on the down triangle next to the **Add Data** button () and selecting **Add Basemap** choose *Imagery with Labels*. Zoom to the NCTC



Click on the Identify button () then click anywhere near the NCTC to pull up the Identify table. On the **Identify window** click the drop down and select *NCTC_Lidar_1m_DEM*.



Now open-up Google Earth on your computer. In the **Search:** enter “National Conservation Training Center” and click on the suggestion then on **Search.**



After zooming to the NCTC elevation of the area is provided (in the image above) in US Standard Feet with a value of 416. Now, back in ArcGIS, use the identify tool on the same area and note the value which should be close to 129. Since 416 is nearly 3.28 times larger than 129, we know that the elevation values are meters.