Handling Raster Data for Hydrologic Applications

Prepared by
Venkatesh Merwade
Lyles School of Civil Engineering, Purdue University
vmerwade@purdue.edu

January 2019

Objective

The objective of this exercise is to learn how to handle raster data in ArcGIS and understand its properties such as data type, coordinate system, horizontal resolution and vertical accuracy in hydrologic applications.

Learning outcomes

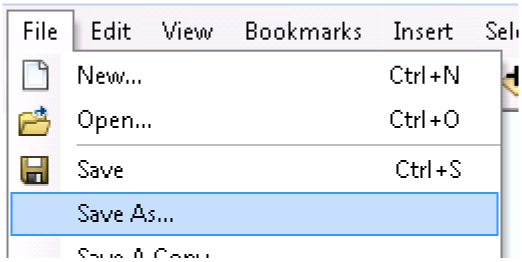
1) Adding and visualizing raster data in ArcGIS
2) Projecting raster data in ArcGIS
3) Clipping a raster data for a given watershed or area of interest
4) Extracting values from a raster data set for points, lines and polygons
5) Computing vertical accuracy of a Digital Elevation Models

Input Data

You are provided with one raster dataset, which is the Digital Elevation Model (DEM), and several vector datasets. These data are available on blackboard as lab2.zip inside the lab2 folder. The data are also available at: ftp://ftp.ecn.purdue.edu/vmerwade/download/data/lab2.zip. Unzip the lab2.zip file in your working folder (keep your folder name and structure simple) and you will see two sub-folders named ned and vector. The ned folder contains the National Elevation Dataset (raster layer) for a region in Indiana. The vector folder contains four shapefiles named point, Tippecanoe, Wabash and watersheds. You will learn more about these layers during the tutorial.

Opening and Saving ArcMap Documents

Open ArcMap as a blank new map by going to Programs…. Save the ArcMap document as lab2.mxd by going to File→Save As..



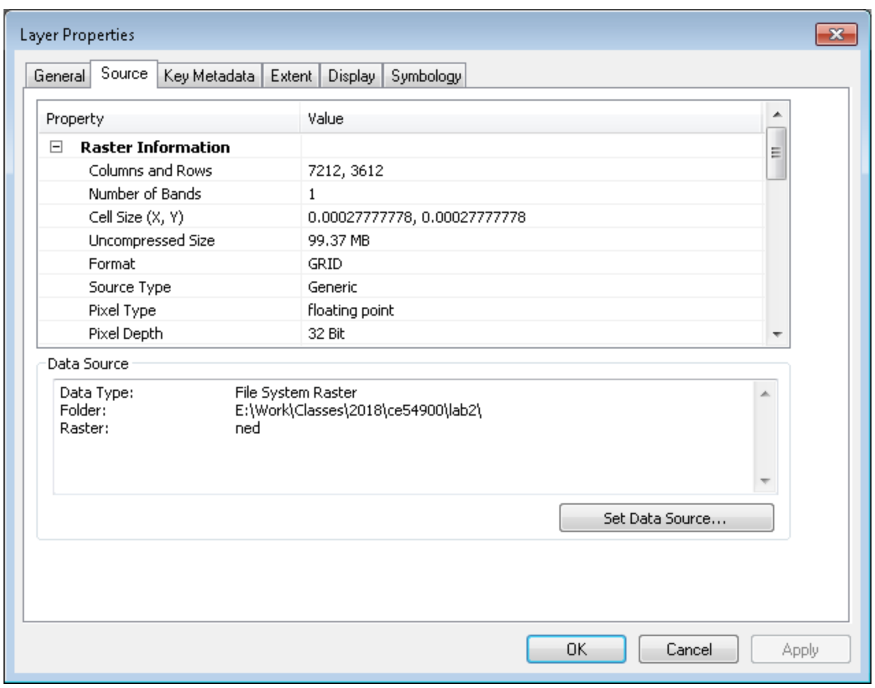
Make sure you save the document at the same location where you have unzipped the input data.

Adding Raster Data to ArcMap Document

Add any type of data to an ArcMap document by either going to File→Add data or by clicking the Add Data button. Add ned data to the map document. If you are asked whether to build pyramids, select Yes. Think about why pyramids are built when you are working with raster datasets. If needed, use ArcGIS Help or any online resource.

Exploring Properties of Raster Data

Lets explore the properties of the elevation data. Right click on ned, and select Properties. In the Layer Properties window, select the Source tab.



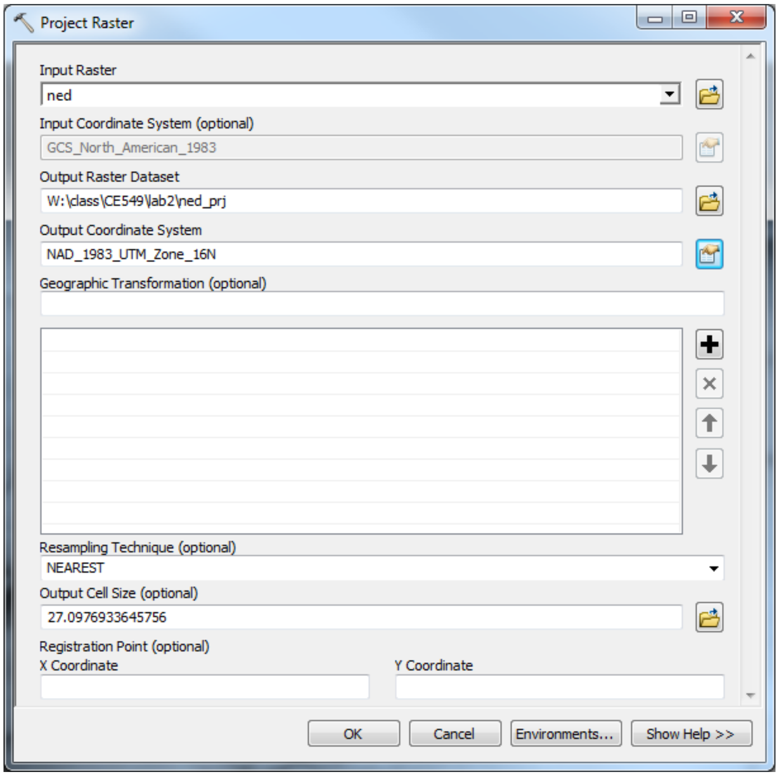
A raster is organized as a matrix of square cells with each cell having a single value. In the case of the elevation data, each cell will have an elevation value to represent the average elevation for that cell. It is very important to know the size of this cell so we know the "horizontal resolution" of a given dataset. The horizontal resolution of a raster is typically referred by using multiple names, including grid size, cell size, or simply resolution. Cell size (x,y) gives the spatial resolution of the data. Because the ned data is not projected, the cell size that you see here is in angular units (degrees). This is one way to tell whether your data is projected to a coordinate system or not. In geographic coordinates, the data will have x,y cell size in degrees. The size of the matrix (number of rows and columns) and the cell size defines the geographic extent of the dataset. Note the cell size of the given data in angular units (arc seconds). Do you know how to convert arc seconds to meters. What will be the approximate size of this DEM in meters.

The next key property in a raster dataset is the data type, which is provided as Pixel Type in the layer properties window. The pixel type for the ned dataset is floating point. What is a floating point data type? What other datatypes can be associated with a raster layer?

The extent gives the (x,y) coordinates of the four corners of the raster dataset. Spatial reference gives the information of the geographic/projected coordinate system for the dataset. Finally, the statistics gives some basic statistics (mean, minimum and maximum values) of the raster dataset.

Projecting Raster Data

Assuming you now know the difference between geographic coordinates and projected coordinates, lets assign some projected coordinates to the ned dataset, and explore the changes to the dataset. If you do not know the difference between geographic and projected coordinates, there is some reading provided on blackboard. Alternatively, you can use ArcGIS help or any online resource to understand this difference. To project any dataset, you will use the Arc Toolbox . Click on the Arc Toolbox. In the Arc Toolbox, select Data Management Tools→Projections and Transformations→Raster→Project Raster



In the Project Raster window, select ned for Input Raster. Name the output raster as ned_prj, and save it at the same location where your other information for this lab is stored. For the output coordinate system, click the button next to it, and select Projected Coordinate System→UTM→NAD 1983→ NAD_1983_Zone_16_N. Change the output cell size to 30 for both x and y, leave all the other default options unchanged, and Click OK. Why are we using NAD 1983 UTM Zone 16 for this data. Remember, Tippecanoe County is located in Indiana. Explore UTM zones within the context of projected coordinate system.

Once ned_prj is added to the map document, open its properties window, and look at the source tab to see what has changed. What changes do you see in the properties of ned_prj compared to ned? How did the coordinates for your data change after doing the projection? What do some of the new terms such as false easting, false northing, central meridian, etc. associated with the projected spatial reference mean. Why do you think this is useful?

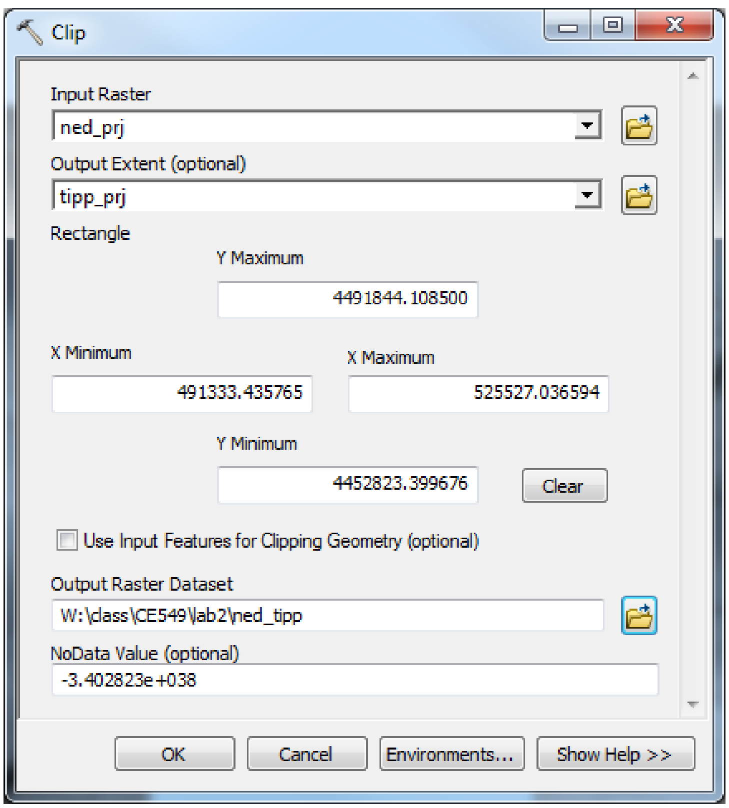
Clipping Raster Data

Most often the data we get is in square or rectangular tile (tiles). When we get multiple tiles, we have to mosaic the tiles to create a single raster (this should be done before projecting). In this case, we have a single tile so we do not have to mosaic. Most often even with or without mosaicing, the spatial extent of a DEM or any other dataset is greater than our area of interest. In such case, to avoid the additional computational burden from extra data, we clip the data to match our area of interest. We call this process as “clipping” or intersection.

Using the Add Data button, browse to the vector folder and add Tippecanoe.shp to the map document. This shapefile gives you the boundary of the Tippecanoe County. Note the coordinate system of the Tipppecanoe boundary shapefile.

Project the Tippecanoe shapefile to the same coordinate system as ned_prj. For projecting ned, you used Raster→Project Raster in ArcToolbox. In this case, you will use Projections and Transformations→Project. Remember Tippecanoe boundary is a vector feature layer and ned is a raster layer. Name the projected feature as **tipp_prj.shp** to save it as shapefile in the same vector folder.

We are interested in getting the elevation data just for the Tippecanoe County. First we will clip the DEM to match with the county boundary. To clip ned_prj, you will be using the Arc Toolbox. Go to Data Management Tools→Raster→Raster Processing→Clip. In the Clip window, select ned_prj as the input and tipp_prj.shp for the output extent. Check the box for using input features for clipping the geometry. Name the output raster as **ned_tipp**, and click OK. For some reason, if Clip does not work for you, try the Extraction tool in Spatial Analyst.

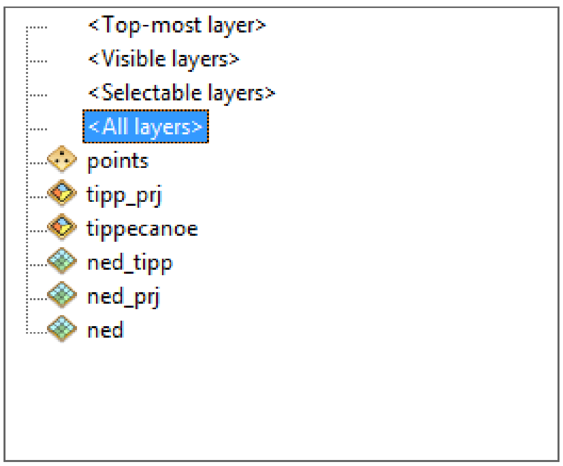


Note the maximum, minimum and average elevation in the Tippecanoe County? (Hint: look at properties)

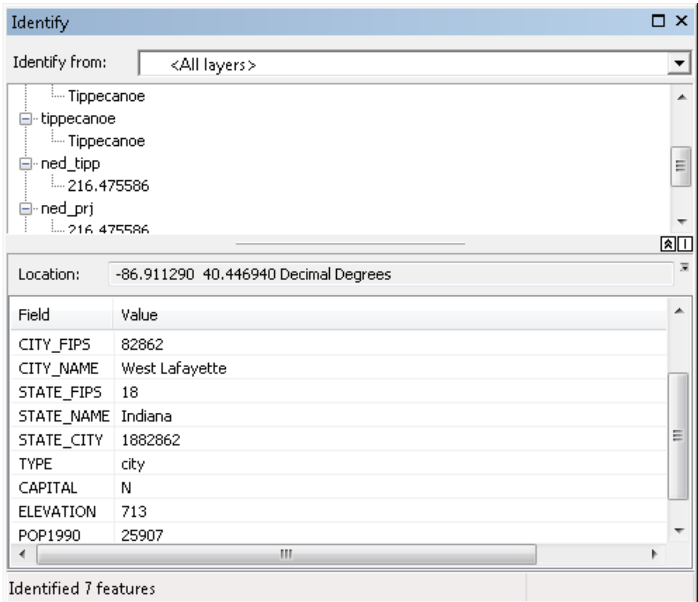
Extracting values from a raster data set for points, lines and polygons

Most often, hydrologists are interested in extracting information for a given location or area from a raster. To extract information at one or multiple points, you can use Arc Toolbox. Add points.shp to your map document from the data folder. Again check if this dataset is projected. If not, go ahead and project it to match its projection with other projected datasets. Name the new point shapfile as **point_prj.shp**.

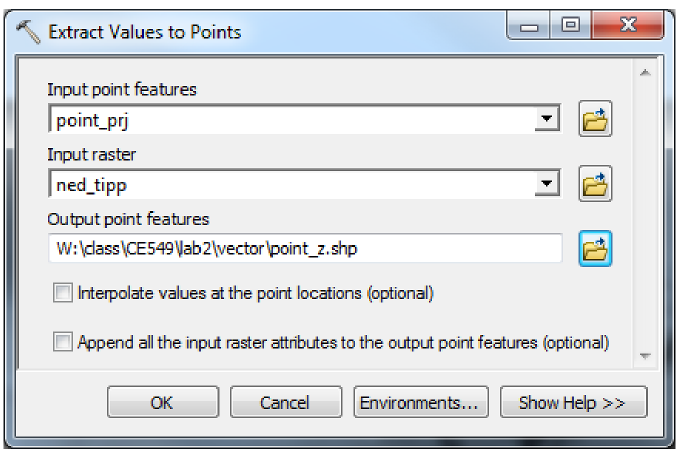
The point shapefile contains the location of Lafayette and West Lafayette in the Tippecanoe County in the form of points. Open the attribute table of points and look at its attributes. We would like to know the elevation at these points. One way to do this is use the identifier button, and click on one of the points. Initially you may only see the attributes of the point shapefile, and not the elevation. Change the settings of the identify window to see the information for all layers (shown below) and hit the identify button again.



You will then see the elevation at this point next to ned_tipp in the identify window as shown below.



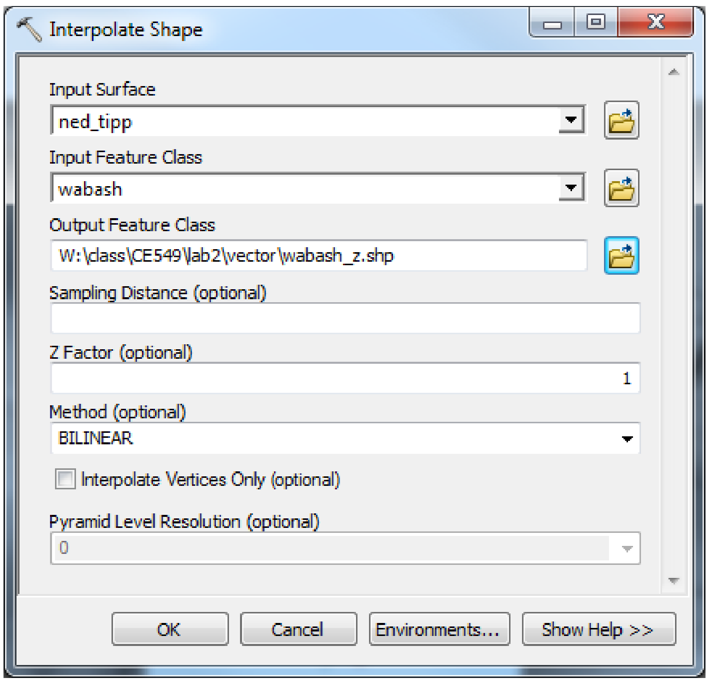
This is a manual way of extracting information for a point from a raster. When you have multiple points, it is tedious to click on each point to get the elevation from the underlying DEM.A better way is to use the ArcToolbox. In Arc Toolbox, select Spatial Analyst Tools→Extraction→Extract Values to Points. In the next window, provide the **point_prj.shp** as input features, ned_tipp as input raster and save the output features in points_z.shp. Click OK.



After the process is complete, points_z.shp will be added to the map document. Open the attribute table of points_z.shp. You will see the extracted values in a field named RASTERVALU. Note the elevation (in meter) of points corresponding to Lafayette and West Lafayette in the DEM.

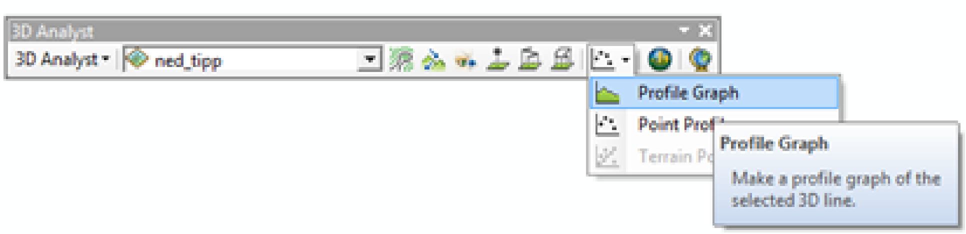
What you just did was extracting elevation for points. In hydrologic applications, elevations also play role in defining the profile of river bed, river cross-sections, river banks and other lateral features. Most lines are 2D, and when you associate elevations with a 2D line, it becomes a 3D line. Add Wabash.shp from the vector folder to your map document. This shapefile describes a short reach of the Wabash river going through Tippecanoe County. Is Wabash.shp projected? Open the attribute table for Wabash.shp to look at its Shape attribute. It should say polyline.

We will now convert the Wabash 2D line into a 3D line. This can be accomplished by using the interpolate shape tool. In the Arc Toolbox, select 3D Analyst Tools→Functional Surface→Interpolate Shape. Use ned_tipp as the input surface, Wabash.shp as input features, and name the output features as **Wabash_z.shp**. Leave the other default options unchanged, and click OK

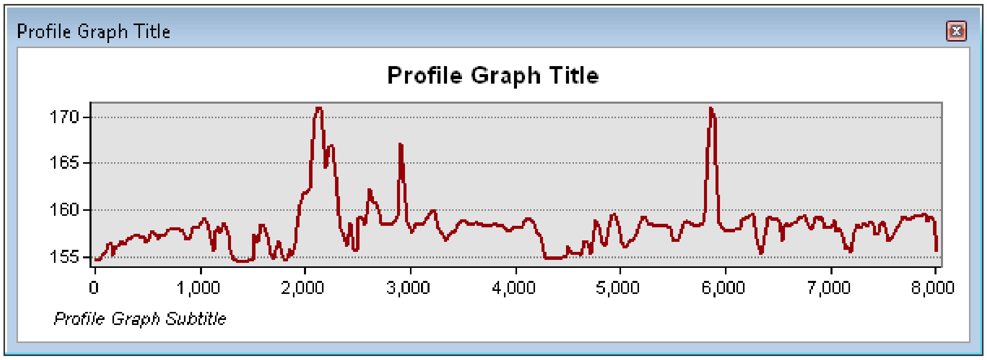


After the process is complete, you will see a new shapefile Wabash_z.shp added to the map document. Open the attribute table of Wabash_z.shp, and look at its shape field. How is it different compared to the shape field in Wabash.shp? What do you think is the difference between a "Polyline" shape and a "PolylineZ" or "PolylineMZ" shape.

Select the line from Wabash_z.shp, and use the profile graph tool in 3D analyst (Make sure 3D Analyst Tooblar is added and the 3D Analyst extension is enabled) to see that the elevations from the DEM are now transferred to the line in Wabash_z.shp to create the river profile (3D line).



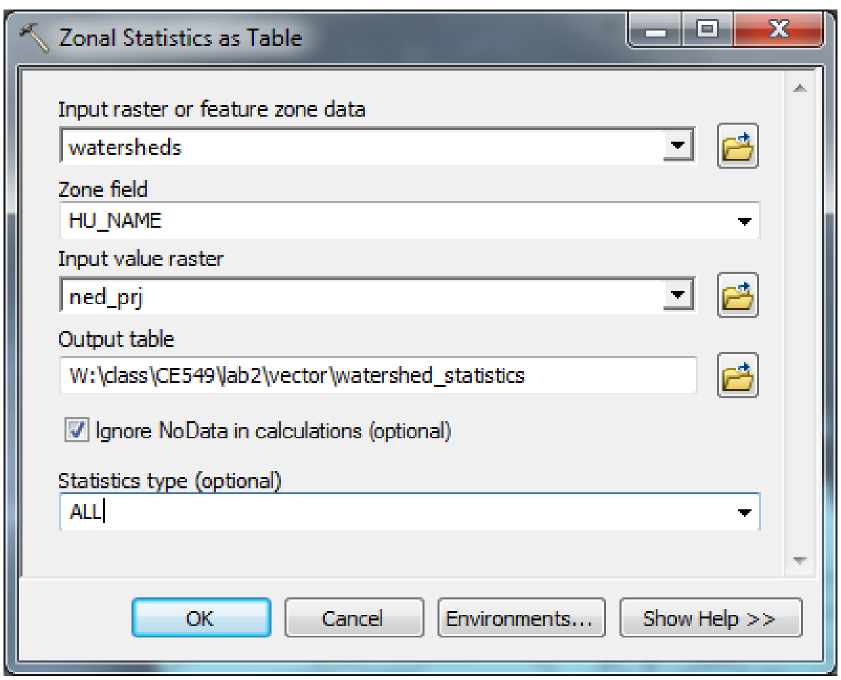
The profile graph tool will plot the elevation profile of the just created Wabash_z line as shown below.



You can follow the same procedure (using Interpolate Shape tool) to convert points, lines and polygons into surfaces.

Now you know how to extract values from a surface to a line or point. In the case of polygons, hydrologists are generally interested in computing average properties for different variables such as watershed slope, average watershed elevation, and average curve number. This task can be accomplished by using the zonal statistics tool. Add watersheds.shp to your map document. Is this layer projected? If not, you should know the drill by now.

What we are going to do now is to compute average elevation for each feature in the watershed layer by using a DEM. In the Arc toolbox, select Spatial Analyst Tools→Zonal→Zonal Statistics as Table.



Input watersheds.shp for input features. Zone field is the field that has some sort of unique identifier for each polygon so that you can use this identifier to link the computed values to the polygons in the input features. In this case, HU_NAME is unique for each polygon. The input raster is ned_prj, and save the output table as watershed_statistics in your working folder. In the statistics type box, you can pick the statistic that you are interested in, or you can compute them all. In this case, we will compute all. Click OK.

The data table that you just created is not linked to the shapefile. Link this table to the shapefile, and create a map that shows different colors for all watershed based on the average elevation. <Hint: use the Join property of the watersheds layer to join the watersheds features with the watersheds_statistics table. Use HU_NAME as the common identifier>. What is the average elevation of the following watersheds: Indian Creek, Buck Creek and Sugar Creek-Little Sugar Creek in Tippecanoe County?

Computing vertical accuracy of digital elevation models

A digital elevation model gives elevation values for each cell. While a finer resolution (small cell size) DEM is desirable, it is more important to have a DEM that is also accurate. That means the elevation that you get for each cell is closer to the ground truth as much as possible. This property is generally referred to as vertical accuracy. If a DEM has a vertical accuracy of 0.5m, the average difference between the DEM and the actual topography is around 0.5m. The more data you have for actual topography, the more confidence you have on the vertical accuracy. In this exercise, you will only learn the process of calculating the vertical accuracy by using the two points we have for Lafayette and West Lafayette in the points shapefile.

The process is relatively straightforward. You need to have points where you have observed ground elevations. In our case, that information is stored in the ELEVATION field (values are in feet) in the point feature class. Next, you need to extract the elevation values from the DEM for these locations. This is already done when you created the point_z shapefile. All you need to do is compute the root mean square error (RMSE) between the two data series, and this will give you the vertical accuracy of the DEM. You can find the equation for RMSE online.

Congratulations!! You successfully finished this tutorial!!