CE 549 Lab 1 - Linking Streamflow Data to a Gauging Station

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

January 2019

Objective

The objective of this session is to learn how to handle vector data (e.g., point) in ArcGIS and link time series to geographic features.

Learning outcomes

1) Creating ArcGIS geodatabase to store geographic and time series data
2) Creating and visualizing vector data (points) in ArcGIS
3) Creating and visualizing Time Enabled Layer in ArcGIS

Input Data

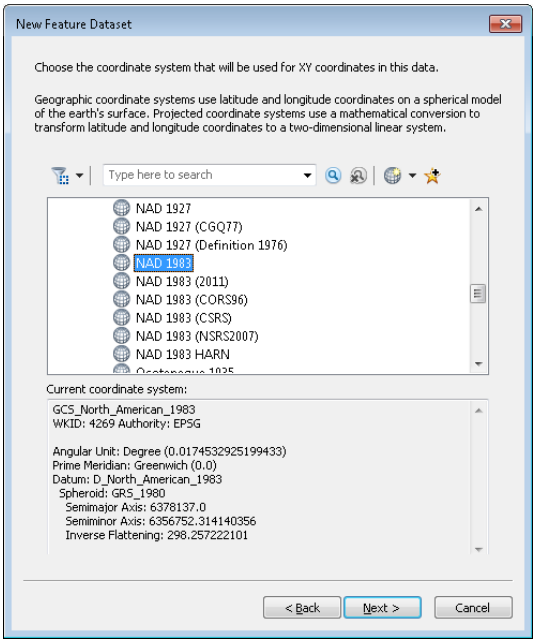
The main input data you need for this exercise is an Excel file with time series information of streamflow. It is provided to you on blackboard as Lab1. Copy all the data in a folder in your working drive (not local drive in the computer lab). The data are also available at:
ftp://ftp.ecn.purdue.edu/vmerwade/download/data/lab1.zip

The data provided to you is the time series data for a streamflow gauging station on the Wabash River in Lafayette (Station Number = 03335500.) The zipped folder contains two files: one is tab separated text file, which is the data as you get from the United States Geological Survey (USSG), and the other file is a comma delimited file which is pre-processed to be used in this exercise. Streamflow data from the USGS can be obtained for any gauge in the US from this link: https://waterdata.usgs.gov/nwis/sw. USGS provides various tutorials that you can use to get the data. The link to the tutorials page is:
https://help.waterdata.usgs.gov/tutorials/surface-water-data

Creating a Geodatabase to Store the data at the Wabash River at Lafayette gaging station

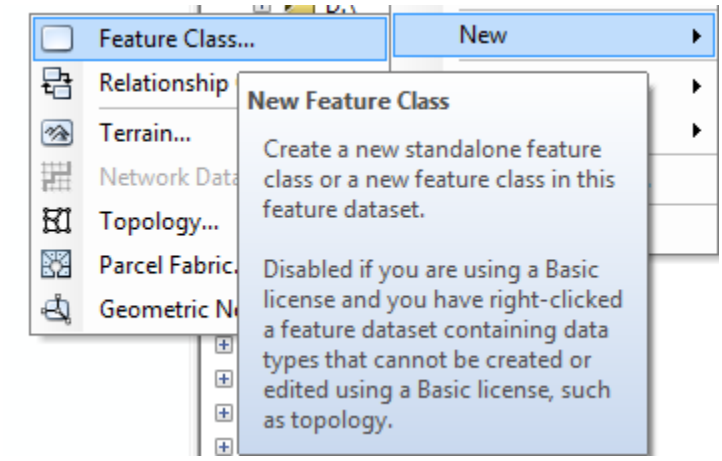
Here we are going to create a point to represent the location of the gauging station for which observations are available. Once the location is created, time series data will be linked to this point. The first step is to create an empty geodatabase where we can store this point and the time series.

Open ArcMap, and **save** it as lab1 arcmap file (with .mxd extension) in your working directory. Remember to keep your directory structure simple without any numbers and special characters. In the Catalog window (if the catalog window is not open in ArcMap, click on the catalog window button ). In your working directory, create an empty personal geodatabase named "lab1." In lab1.mdb, create a feature dataset named Data. Right click on lab1.mdb, and **Select** *New→Feature Dataset*. **Name** the feature dataset as "Data". **Assign** North American Datum 1983 geographic coordinate system (NAD 1983) to the Data feature dataset by selecting Geographic Coordinate Systems→North America→NAD1983). This information comes from the gauging station location map on the USGS website (link: https://waterdata.usgs.gov/nwis/nwismap/?site_no=03335500&agency_cd=USGS).

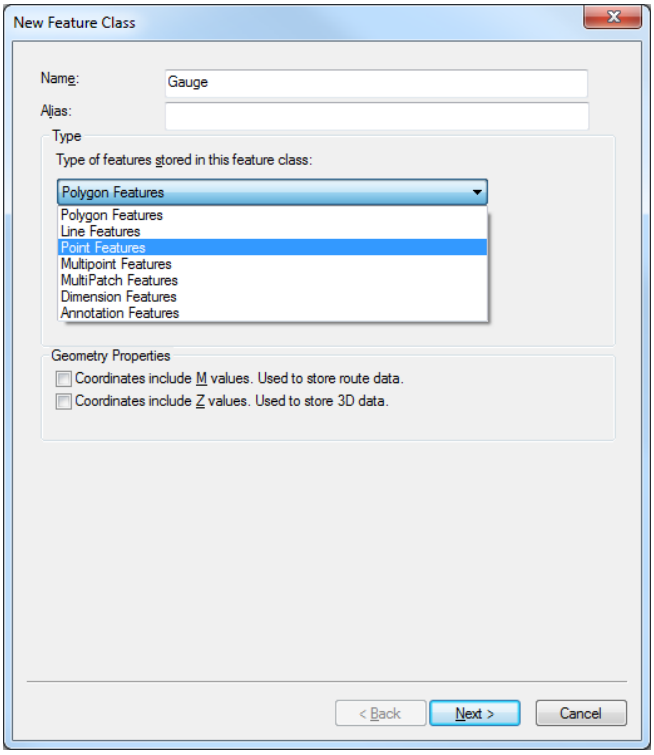


For vertical coordinates, use NAVD88 by selecting North America→NAVD 1988. Leave the default x,y tolerance unchanged, and finish creating the feature dataset.

Next, **right click** on Data feature dataset you just created, and create a new feature class as shown below.

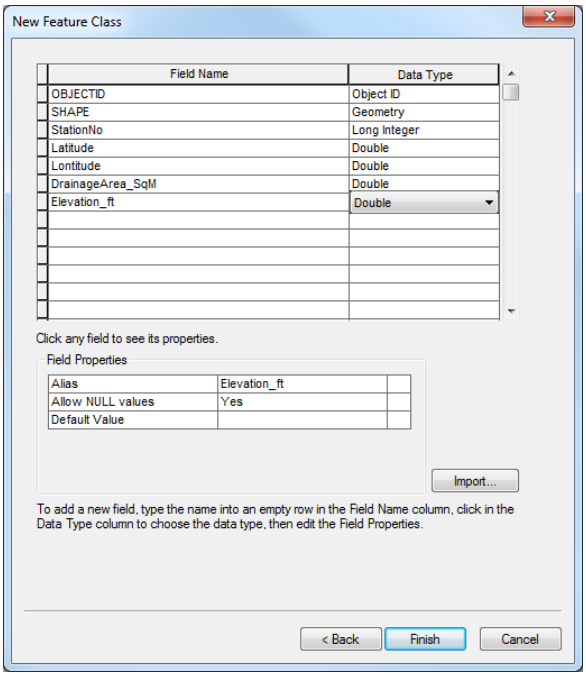


Name the new feature class as "Gauge" and make it a point type feature class as shown below:



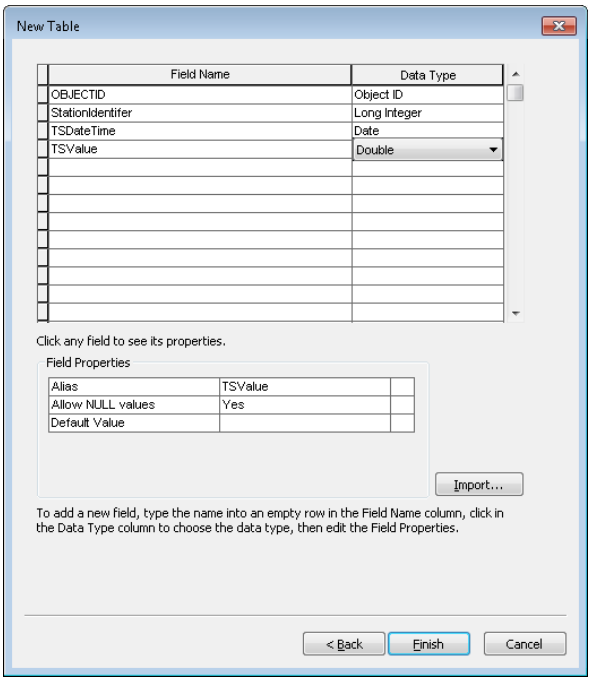
Leave M values and Z values unchecked for this feature class. **Click** *Next*.

In the next window create some new fields to store information associated with a gauge such as its station number, location, name, etc. So create the following fields with the corresponding data types (Most names are self-explanatory. _SqM and _ft after drainage area and elevation, means the area and elevation values are reported in reported in square miles and feet, respectively):



Click *Finish*. You will see that the Gauge feature class, which is empty, is now added to the map.

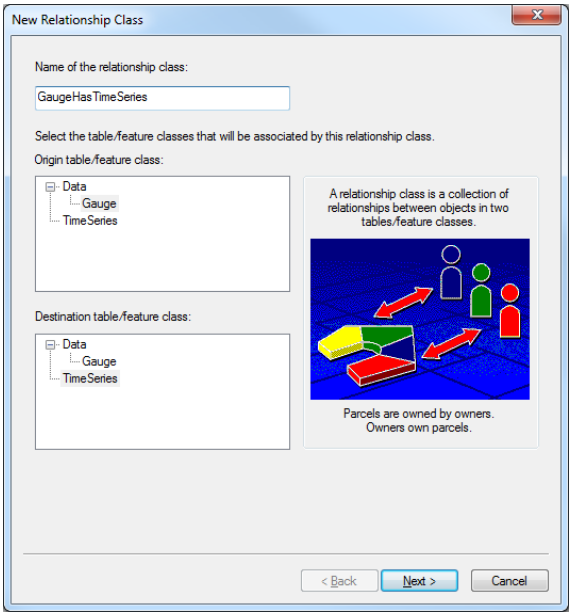
Next, **create** an empty table in lab1.mdb, and name it as "TimeSeries". Follow the same steps as above, but instead of creating a feature class, create a table by right-clicking on Lab1.mdb (not on Data feature dataset). You will also realize that when we create a table, we do not specify its type (point, line or polygon), or coordinate system because we do not store any geographic data in a table. How do you think a table is similar or different than a feature class? Define some fields for TimeSeries table as shown below:



StationIdentifier is an identifier for the location of the gauging station, TSDateTime is the time stamp for the observation (when the observation was made or recorded), and TSValue is the measured or observed value. "TS" in TSDateTime and TSValue stands for "Time Series". It is not necessary that you give these exact names to the fields, but it is good to follow some standardized naming procedure so you will have consistency among all your datasets.

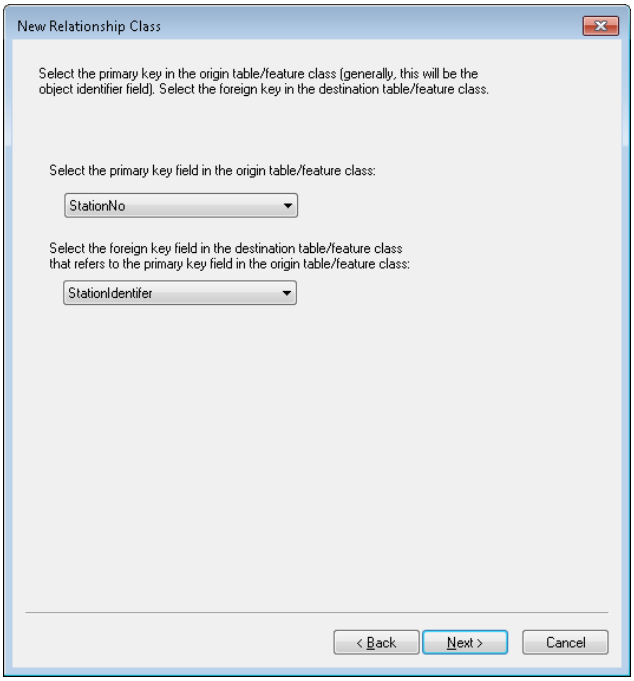
Click Finish. What we have done here is first create an empty point feature class to store the point and then an empty table to store the time series. One last thing before we are done with creating this dataset is to define a relationship class that will link the Gauge geographic feature with TimeSeries data. In this exercise, we are dealing with one layer and one table, but in real world you may have multiple tables and multiple feature classes so again as a good practice, we will name the relationship class in such a way that it will tell which two datasets are related.

Right click on lab1.*mdb,* and **create** a new relationship class named "GaugeHasTimeSeries". **Select** Gauge as the origin table and TimeSeries as destination table as shown below:

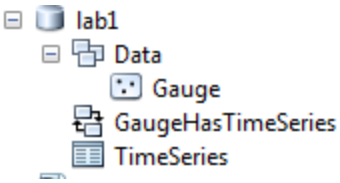


In the next few windows, **select** the following options: simple peer-to-peer relationship, no messages are propagated, one-to-many relationship (one point will have many measured values), no attributes for the new relationship class.

Finally, use StaionNo in Gauge and StationIdentifier as the primary key to link these two tables. Note that the fields do not need to have the same name, but they should be of same type and contain identical information. In this case, both fields will contain the station number for the gauging station.

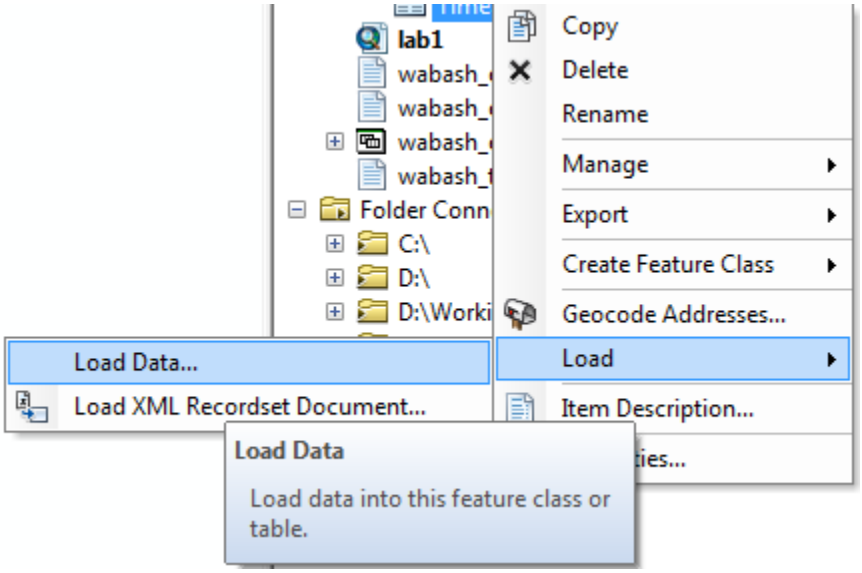


Click *Next* and then *Finish*. Your final geodatabase in Catalog view should look something like below if you use the same names as suggested in this exercise.



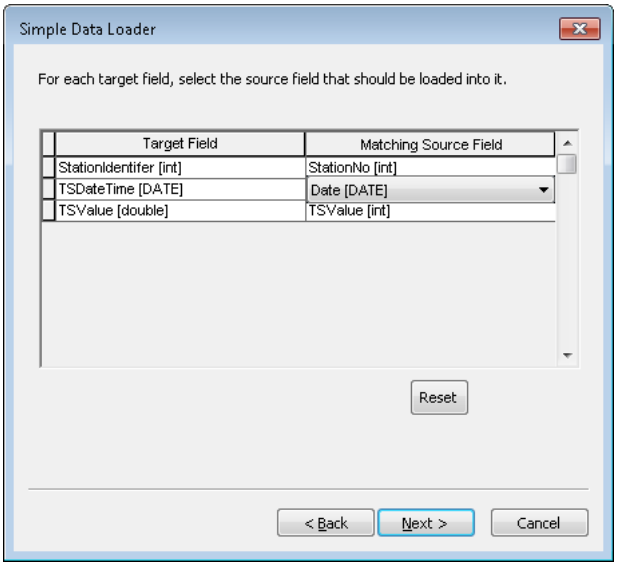
What we have now is an empty database. Our next step is to populate this database. First, we will populate the time series table. The discharge data we need for TimeSeries already exist in the comma separated file so we can load that data into TimeSeries in ArcCatalog.

Right click on *TimeSeries*, and **select** the *Load* data option as shown below.



Browse to the Wabash Data CSV file, **click** *Add* and then **Click** *Next*.

Click next on the subsequent window, and then match the following fields as shown below.



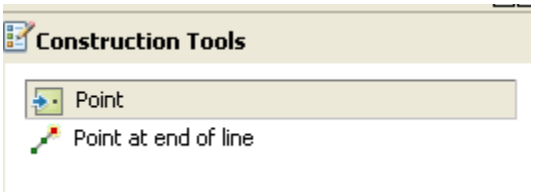
Click *Next* and then *Finish*. If you open the TimeSeries table (right click in ArcMap), you should see 365 average daily streamflow values for the Wabash River in Lafayette in 2018. Now that we have the data in TimeSeries, the next step is to create a point for the gauging station and explore this link through the relationship class that we just created. Very cool! **Close** the ArcCatalog window.

Populating Gauge data in ArcMap

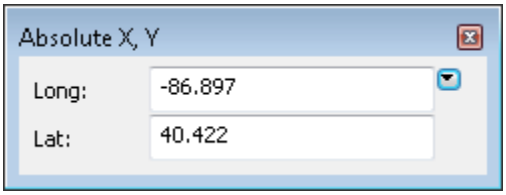
Your map document should already have the Gauge featureclass and TimeSeries table added in the table of contents view. If not, **add** Gauge and TimeSeries to the map document using the Add data button, or by going to the file menu. **Save** the map document.

Open the editor by pushing the editor toolbar button . To create a point for the gauge, **click** on the *Editor* toolbar and **Start** Editing.

Once the edit session is started, an editor window will be added to the map document that will show the features and the construction task. Select the Gauge feature class in the editor window, and the construction tasks associated with this feature will be listed in the Construction tools as shown below.



Select the Point construction tool, and bring the cursor to the map document. Now, to add a point by using the lat-long of the gauging station, **press** *F6*. You will be asked to enter the absolute X and Y for the new point. For X, enter the longitude for 03335500 with a negative sign (western hemisphere) and for Y, enter the latitude for 03335500 (both in decimal degrees), and **press** enter on your keyboard.

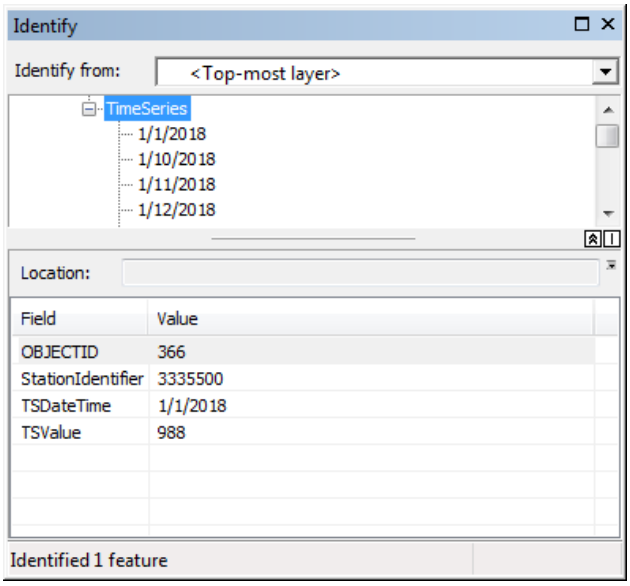


You will see that a new point is created at the specified location on the map document which is stored in Gauge feature class. If you do not see the point, right click on Gauge feature class and select "Zoom To Layer". Now open the attribute table for Gauge (by right-clicking), and populate all the fields using the information for 03335500. Other information that you need is the drainage area (7267 square miles) and the altitude (= 503.84ft). All this information comes from the gauging station location information on the USGS website. Remember your station number in the gauge feature class should match exactly with the station number in the time series table. Close the attribute table, save edits by clicking on editor toolbar, and stop editing. Remember, you can create and edit features only during an edit session. If you make a mistake during the editing process, you can stop editing without saving the edits, and your data will be unaffected from your wrong edits.

What you just did is very cool and also learned several GIS functionalities. We can now identify in GIS where 03335500 is located. To make this location relevant, you can add other GIS layers such as counties, state boundaries and stream lines which you can find in ArcGIS online. To add data from ArcGIS online, click the add button and select Add Data from ArcGIS Online… as shown below. Adding this may take a while so you can be patient or do this after the lab!

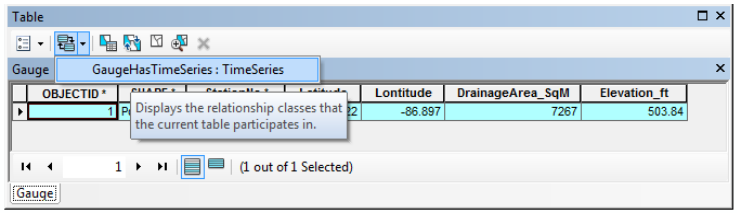
OK, now that we are confident about the location of our gauging station, the next step is to see how Gauge and TimeSeries are linked together. First, right click on TimeSeries table. Go to Properties→Display, and change the display expression from StationIdentifier to TSDateTime

Use the identifier button , and click on the gauge. You will see the identify window, that will display all information about gauge attributes as shown below. You can click on a date, and you will see what the streamflow value on that day is for this point.

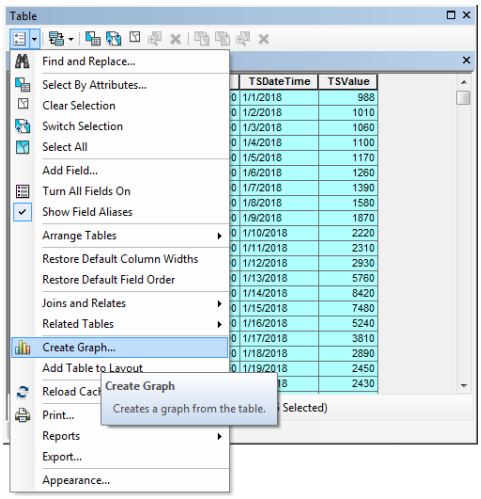


TimeSeries Plotting

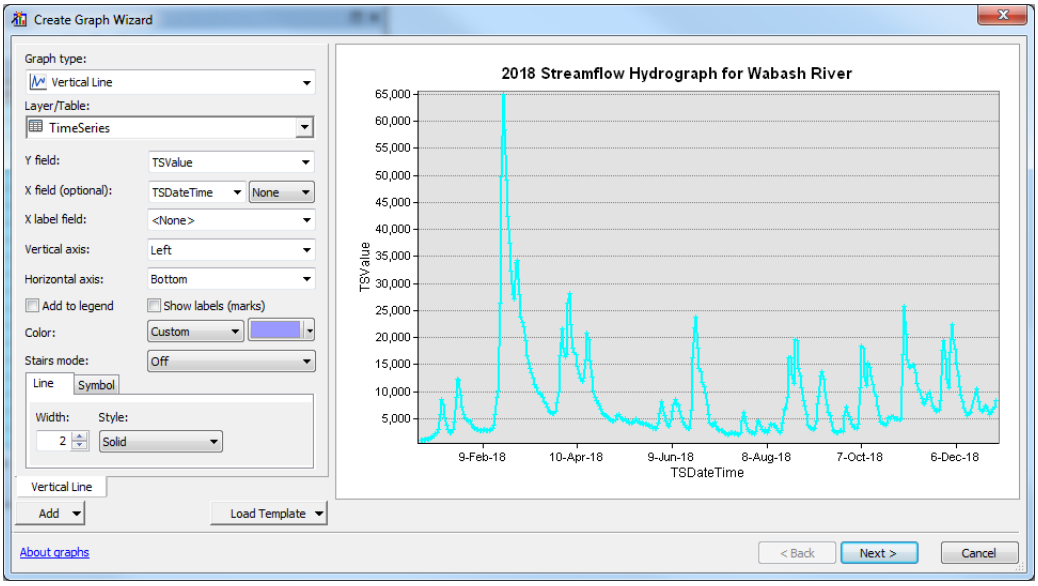
You can even plot the data associated with a gauging point by selecting the point, and getting access to its related time series. Open the attribute table of the Gauge feature class, and select a feature (we have only one). To select the feature, click on the beginning triangle button in the row. Next, click on the Related Table button in the attribute table menu, and click on the GageHasTimeSeries: TimeSeries button as shown below



This will open the TimeSeries table with the time series values related to the selected point as selected rows. Now you can plot the selected rows in a graph as hydrograph by clicking on Table Options→Create Graph as shown below.



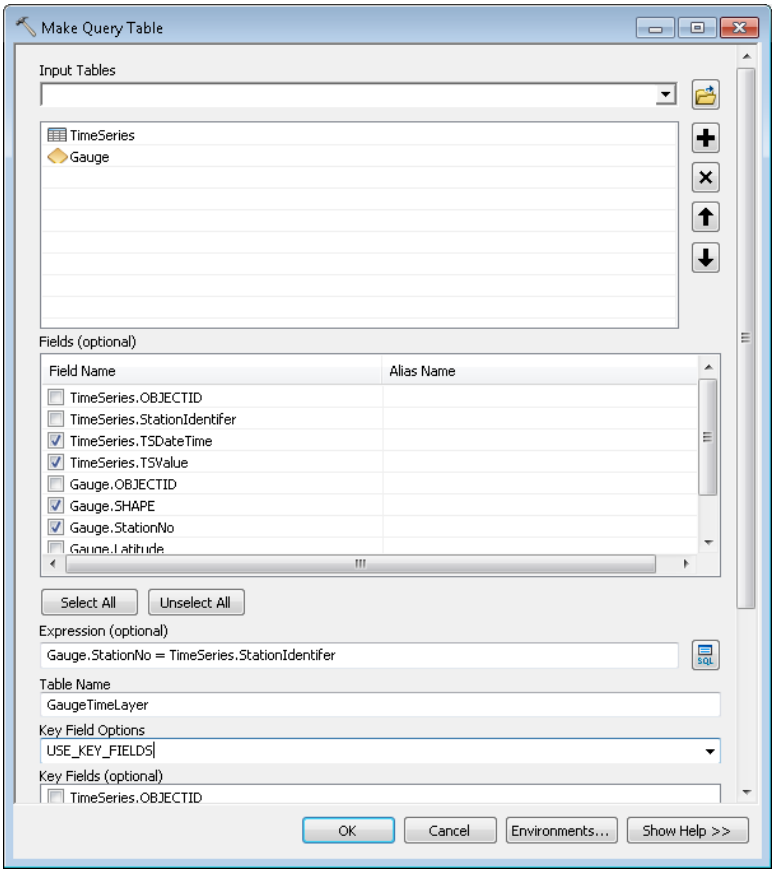
In the create graph wizard, choose the Y field to be TSValue and X field to be TSDateTime.



Change other variables as needed, add labels, title, etc, and create the graph. What you have learned so far is to bring the time series data into GIS, link it to a geographic feature and plot a graph! If you think this is cool, what we are going to do next will be super cool!! We are going to visualize the data as it changes in space and time.

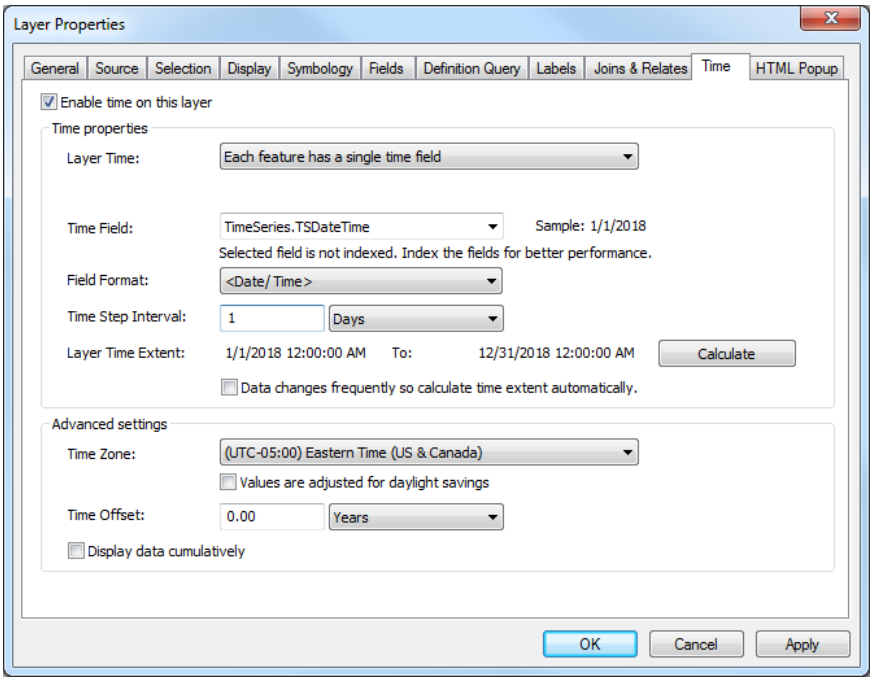
Creating Time Enabled Layer in ArcMap

By using the link between GIS features and related time series, we can also create a time enabled layer in ArcMap and produce animations. To create a time enabled gauge layer, lets first join the Gauge and TimeSeries tables through a query table. Open ArcToolbox, and select Data Management Tools→Layers and Table Views→Make Query Table. Add Gauge and TimeSeries as the input tables. For fields, make sure to select the StationNo and SHAPE in the gauge feature class, and TSDateTime and TSValue in the TimeSeries table. For the expression, use the SQL query button to equate the related fields in Gauge and TimeSeries (StationNo and StationIdentifier in our case as shown below. Make sure there are no single/double quotes in your expression). Name the output table as GaugeTimeLayer, leave the other defaults unchanged, and click OK.

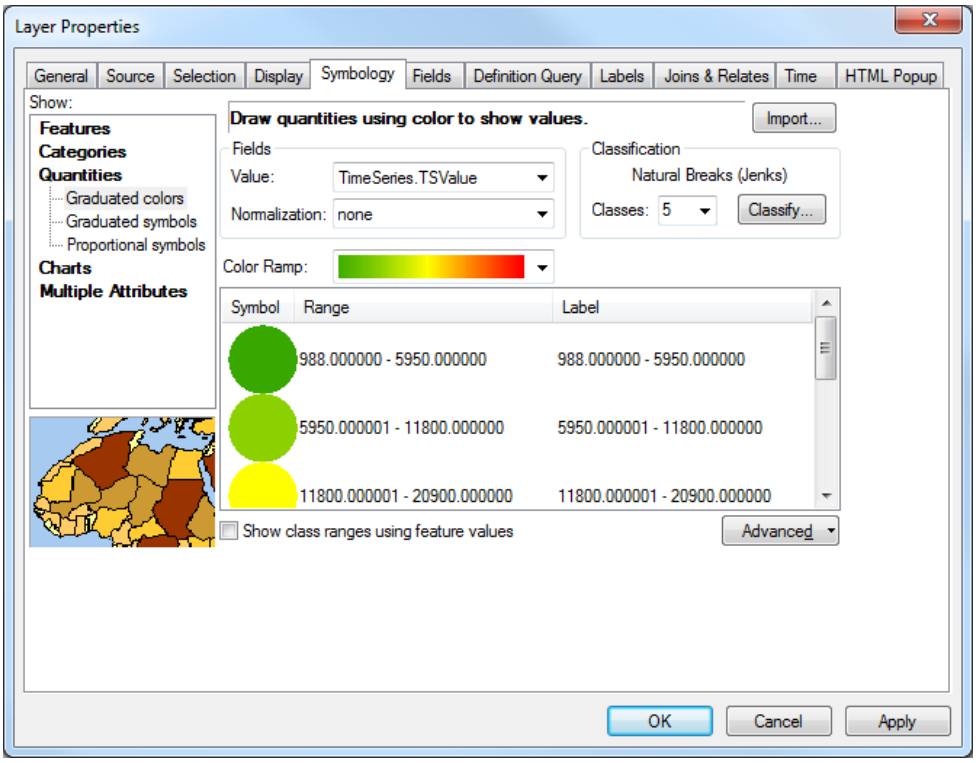


This will create a new temporary feature class named GaugeTimeLayer, and add that to the map document. If you open the attribute table of this new layer, you will see that there is a point created corresponding to each related time series record for a gauge point.

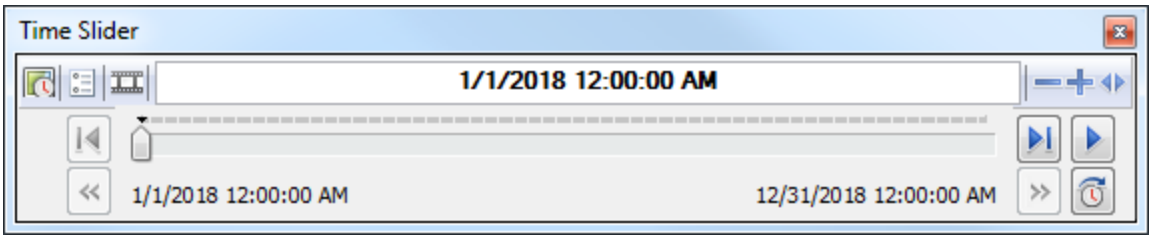
Next, goto GaugeTimeLayer properties, and select the Time tab. Once you see the properties click on the Time tab. Check the Enable time on this layer box, make sure the right time field is selected, and change the time step interval accordingly (we have daily streamflow data so selecting one or more days is OK) as shown below. Press the Apply button and then press OK.



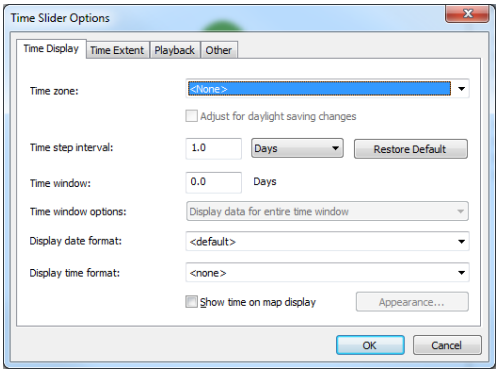
Next, select the Symbology Tab, and modify the symbology to have different color or shape for the point based on TSValue by using graduate colors or symbols.

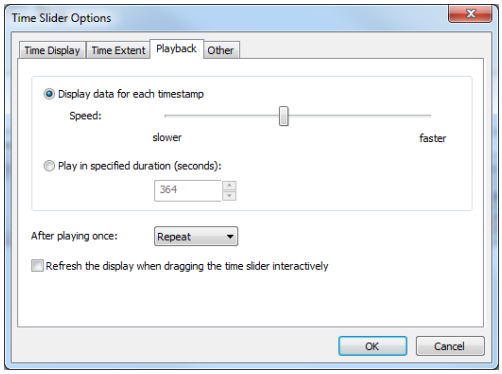


You can also label the feature to show the TSValue. That way you will see both the values and symbol as we navigate through time using the new layer. Once this is all done, the time slider tool  on the Tools toolbar will be activated. Click on that to see the time slider bar as shown below.

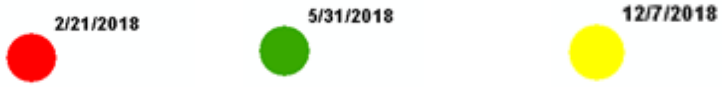


Click on the Options button on the time slider bar, and make sure the time step interval matches with the data, and change the playback speed appropriately so that the playback is not too slow or fast as shown below.





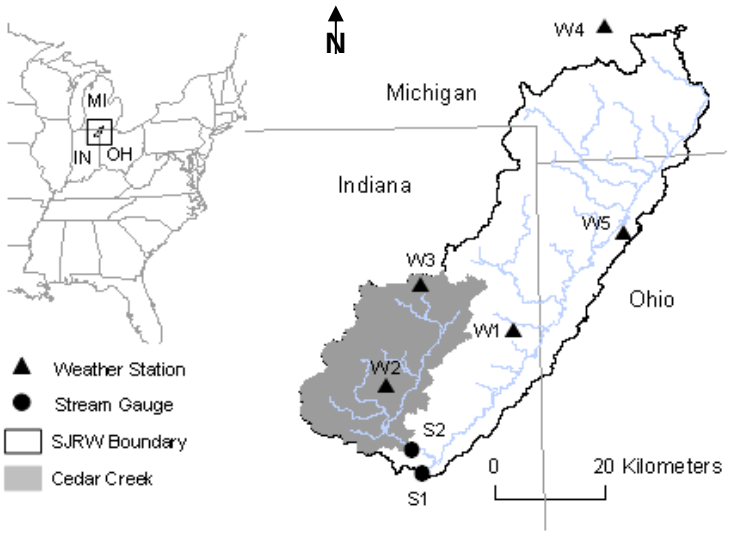
Now hit the play button on the slider bar, and see how the symbology of the point changes based on streamflow value for different days as shown below.



What you just learned is a powerful functionality in GIS to link geospatial and temporal information together that can be used for data analyses, visualization and animation!

HOMEWORK (Due on 01/18/2019 by 5:00 PM)

Below you see a map of St. Joseph River Watershed (SJRW) in northeast Indiana. We are interested in downloading data for streamflow station (S2) which serves as the outlet for Cedar Creek (one of SJRW sub-watersheds). S2 represents Cedar Creek near Cedarville, Indiana station with 04180000 station number as its unique identifier.



- Using the NWIS web interface download daily data for S2 from 01/01/2018 – 12/31/2018. (Hint: select the station by using "Site Name Identifier", and search the site name by choosing the "match any part" option)
- Set the desired data range (check date format), and choose the output option as "Tab-separated data" save to file. (you can also display the tab separated data in a browser and then save it).
- Also get the site description by using the "Location Map" option, and report the following information for 04180000: HUC Number, latitude (decimal degrees), longitude (decimal degrees), geographic projection, drainage area and datum above sea level including reference.
- Open the tab-separated text file into Excel and save the file as comma separated file in the same format as Wabash comma separated file.
- Create a point feature for the Cedar Creek gauging station in the same feature class where you have stored the Wabash gauging station. Populate all the fields existing fields with necessary information for the Cedar Creek Gauge.
- Load the time series data for S2 in the same TimeSeries table where you have data for the Wabash River.

Turn-in the Following (Hard printed copy only)

1. Separate presentation quality plots of daily flows for both 03335500 and 04180000. You will use ArcGIS (not Excel to make these plots). Report the 2018 flow statistics (mean, standard deviation, minimum and maximum) for both these stations.
2. How can you distinguish between the data for Wabash River and Cedar Creek when the data are stored in the same table?
3. If you give your database to another person, can that person tell what the unit is for the values stored in your TSValue field? What improvements (e.g., addition of new fields if any) do you see in the database that you have created for storing streamflow records?

4. How can you store precipitation information in the same TSValue field in the TimeSeries table in your geodatabase? How can you distinguish between precipitation and streamflow values if both are stored in the same field? What additional improvement will you make to the TimeSeries table to distinguish different types (streamflow, temperature, precipitation, etc.) of data values in the same column?

Answer the following questions (do not copy/paste text from ArcGIS help or other online sources):

i. What do you think are the key functionalities of ArcMap, ArcToolbox and ArcCatalog?
ii. What is vector data in GIS?
iii. What is a geodatabase?
iv. What is the difference between a personal geodatabase and file geodatabase?
v. What is a feature dataset?
vi. What is a feature class?
vii. What is a relationship class?
viii. What is the difference between a feature class and a table in ArcMap?
ix. What is a feature in ArcGIS?
x. What is a field in ArcGIS?
xi. How can you access attributes and other properties of a feature class or table in ArcMap?

Answer yes/no to the following? Make sure to think or repeat these steps before answering Yes/No.

a) Can you open ArcMap, ArcCatalog and ArcToolbox in ArcGIS?
b) Can you create and save a new arcmap document to a specified location?
c) Can you create and populate a personal geodatabase in ArcGIS?
d) Can you assign spatial coordinates to GIS data?
e) Can you add and populate fields in any feature class or table?
f) Can you change the symbology for a vector feature in ArcMap?
g) Can you edit a geographic feature or table row in ArcMap?
h) Can you start, save and stop an edit session in ArcMap?

Email your personal geodatabase for this HW as zip file (yourlastname.zip) to vermerwade@purdue.edu