Analysis and Comparison Between the USGS Standard Method and the ESRI Ready-To-Use Method of Watershed Delineation

Emily Pease

The University of Texas at Austin, Energy and Earth Resources

December 7th, 2018

1 Background and Introduction

Watersheds, also known as catchment areas or drainage basins, are areas bounded by topographic highs in which all precipitation that collects will drain out a common outlet, such as a river, bay, or other body of water. Understanding where watersheds are located and how large or small they are helps hydrologists and decision makers determine basin characteristics, flooding, and water availability in a region. Central Texas experiences extended times of drought interspersed with intense flooding events, increasing the importance of understanding the watersheds and water availability in the region. The United States Geological Survey's (USGS) National Water Information System (NWIS) is the database in which all USGS water data is stored and available online to the public. The USGS installed and maintains hundreds of streamgages located along major and minor rivers in the state to capture real-time water data. Streamgages are located throughout the country and provide essential data to water management who rely on this data daily and especially during emergency situations. There are a higher density of streamgages on rivers prone to flooding or high water level variability typically, but there also exist streamgages on small creeks and steams. In NWIS, each streamgage contains information including the site identification number, location, available period, hydrologic unit code (HUC), drainage area, contributing drainage area, and other parameters including precipitation, discharge, and gage height. Boundaries, or watershed delineations, are drawn in using geographical information systems (GIS) software. The constant data collection, field work, construction, and upkeep of the streamgages provides a large amount of the total USGS Texas Water Science Center funding each year. Because new streamgages are installed each year, it is likely that there will always be a need and financial ability to generate watersheds for each of these new sites.

Software and programs exist that help an analyst create a delineation of a watershed. These include ESRI [2], QGIS [4], and computer programming (i.e. Python), among others. There are benefits and disadvantages to all of these but each produces a similar output. In the Geospatial Sciences and Cyber Innovation branch in the Texas Water Science Center of the USGS, ESRI's Arc Suite is the main GIS software used for geospatial analysis. This has been the default GIS software since the branch was established over a decade ago and has remained the trusted method because the output is assumed accurate and reliable. With ESRI, the team is able to delineate watersheds using the National Elevation Dataset (NED) as the initial raster dataset, and send them off for specialist review to be added to the NWIS database.

The watershed size varies based on the streamgage location and size of the river it is installed on. When larger watersheds are requested, processing takes hours to days to complete using ESRI. Therefore, GSCI has been searching for a more efficient method. A method of consideration for watershed delineation using ArcGIS Pro is the ESRI Ready-to-Use watershed tool. This tool delineates watersheds in a fraction of the time that the USGS workflow requires, which makes watershed delineation a more timely process. Here, the USGS method will be compared against the ESRI Ready-to-Use method of watershed delineation. These two methods produce watershed boundaries using ESRI tools, but the workflows vary.

For the purpose of clarity, the method used by the USGS is referred to as the USGS method, though it uses ESRI watershed processing tools for geoprocessing. The Ready-to-Use watershed tool workflow will be referred to as the ESRI method.

2 Objective and Scope

The current process for watershed delineations is tedious, time-consuming, and computationally expensive. Once the ESRI Ready-to-Use hydrology tools were discovered early in this course, it appeared to be a promising alternative method to the expensive data processing that currently takes place.

The purpose of this class project was to provide a comparison between the current USGS watershed delineation method and the ESRI Ready-to-Use watershed tool to determine how reliable the ESRI tool is and to understand its limitations in the USGS scope of work. Additionally, the goal of this project is to make an informed decision on the ESRI method's suitability as a replacement to the current USGS method.

3 Methodology

3.1 Datasets and Resources

The National Elevation Dataset (NED) 10 meter digital elevation models (DEMs) from 2013 [1] were used as the input raster for watershed processing. This dataset is used by the USGS because it is nationally available and has a high spatial resolution. ArcGIS Pro was the software used for watershed processing to compare the USGS process to the ESRI process. All data for this study was publicly available to download and was stored locally.

3.2 Workflows

3.2.1 USGS WORKFLOW

In the Geospatial Sciences and Cyber Innovation Branch (GSCI) of the USGS, there is a specific procedure that is followed to calculate drainage basin areas to be entered into NWIS [3]. The current process is the following:

1. Locate the newly installed streamgage in ArcGIS Pro [2] and create a "pourpoint" shapefile at its location using the Editor tools

2. Examine the National Elevation Dataset [1] gridded shapefile and identify the correct Digital Elevation Model(s) that cover the desired basin area
3. Mosaic the DEMs (if there are multiple)
4. Reproject the mosaiced DEM to the Texas Centric Albers Equal Area projection (ESPG 3083)
5. Fill sinks
6. Calculate flow direction
7. Calculate flow accumulation
8. Snap the pour point to the flow accumulation raster
9. Create watershed raster
10. Convert waterwshed raster to polygon
11. Convert and export the watershed polygon as a KML and send to surface water specialist for review

This process takes between one hour and 10+ hours to complete based on how large the watershed is and, more importantly, the resolution and number of DEMs that were used in the process.

3.2.2 ESRI WORKFLOW

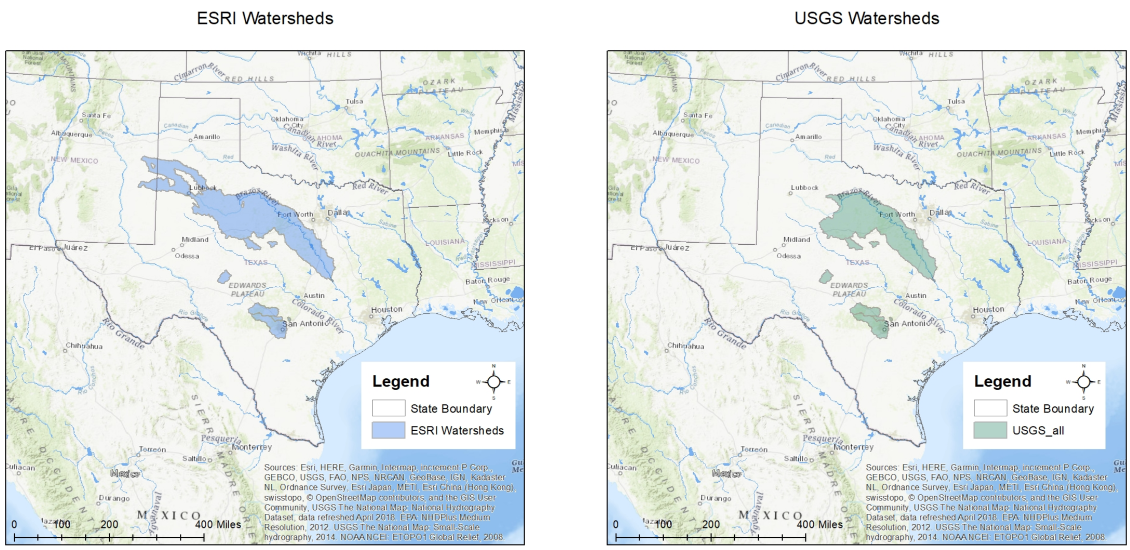
The ESRI Ready-to-Use toolbox contains elevation, hydrology, and network analysis geoprocessing tools. Here, the hydrology toolset's watershed tool was used to calculate basin areas for the ESRI method. This tool works at a much faster rate than completing the process one ESRI tool at a time, as is done in the USGS method. All data and processing tools are black box (hidden on the back end) to allow the user to simply select the watershed tool, input their pour point, and hit "run", without the need to fully understand the computations being done by ESRI. The downside of being black box is that the user doesn't know exactly what the computer is doing and the code can't be edited as needed. Both methods have this aspect to them and it should be noted here that this is not just present in the ESRI method, but is present in ESRI software altogether.

The workflow for the ESRI method is the following:

1. Create a pour point shapefile from a watershed is to be delineated
2. Run watershed tool
3. Save watershed polygon

3.3 Methods

There are essentially two components to this project, one being watershed delineation using the standard USGS method, and the second being a second delineation of the same sites using the ESRI Ready-to-Use watershed tool. These two methods will be compared in order to determine how reliable the ESRI method is in comparison to the USGS method. Because the USGS is the gold standard of hydrologic data and ESRI is the proprietary gold standard for GIS analysis, it will be interesting to see how the two methods watershed outputs compare.



(a) ESRI method watershed polygons. (b) USGS method watershed polygons.

Figure 1: Maps of the ESRI method and USGS method watershed outputs. There are clear differences visually from these two methods that are addressed in Section 4.

Twelve NWIS sites were selected as pour points for watershed in this study. They are NWIS sites 08098450, 08140860, 08180990, 08128030, 08167200, 08168770, 08061548, 08099382, 08178980, 08181725, and 08183978. They represent NWIS sites ranging in basin area located throughout Texas. Two watersheds were processed for each pour point (24 watersheds in total) following the workflows as described in sections 3.2.1 and 3.2.2. Digital elevation models may often have been created after the construction of man-made topographic features and the boundary in these areas won't be captured accurately. Therefore, it is typical for watershed boundaries to require slight editing after they've been processed due to the presence of major highways, dams, or other man-made topographic features built after the DEM was created. However, for the purposes of this study, no editing took place after a watershed boundary was produced in order to measure the area difference from raw outputs.

There were a few details in each workflow that must be addressed. For both methods, the D8 method of calculating flow direction was used (Equation 2). The direction of flow is determined by the direction of steepest descent, or maximum drop, from each cell. This is what the USGS uses and it was selected for both the USGS and ESRI method for consistency. In ESRI's Ready-to-Use watershed tool, the user is given a choice of DEM from best available, 10m, 30m, and 90m. The best available DEM was selected for each watershed in the ESRI method because there exists 10m data for the continental US that is available to ESRI, and 10m DEMs were used for the USGS method.

The percent difference in area for the twelve NWIS sites (24 watersheds) was determined. The area in square miles for each watershed using both methods was calculated in ESRI and compared against each other using Equation 1.

$$\%_{diff} = \frac{|Area_{ESRI} - Area_{USGS}|}{\frac{(Area_{ESRI} + Area_{USGS})}{2}} * 100\% \tag{1}$$

$$drop_{max} = \frac{\Delta Z}{distance} * 100 \tag{2}$$

Because this study's primary purpose was to determine the reliability of the ESRI method in comparison to the USGS method, a statistically significant number of watersheds were not processed. Twelve watersheds proved to be an adequate number to determine whether the current method could be replaced by the ESRI method.

Site_ID	USGS_Area (sq mi)	ESRI_Area (sq mi)	PercentDiff
8098450	17114.883	27024.59	44.902
8140860	450.06	448.976	0.241
8180990	38.478	39.124	1.664
8128030	446.106	437.079	2.044
8167200	1102.669	1102.872	0.018
8168770	29.707	30.262	1.851
8061548	0.308	0.047	147.53
8099382	143.663	143.668	0.003
8178980	472.186	472.217	0.007
8181725	1719.834	1720.062	0.013
8183978	162.265	162.497	0.143

Figure 2: Table showing percent differences for the twelve USGS watersheds and ESRI watersheds

4 Results and Discussion

Twelve watersheds were delineated using two methods and their areas were compared (Figures 1a and 1b). The percent differences for the twelve NWIS sites were calculated and they can be found in Figure 2. In Figure 2, the site ID is the NWIS unique identifier given to the streamgage by the USGS. The USGS and ESRI area columns contain the watershed areas in square miles using both methods. Ten of the twelve watersheds in this study produced a small percent difference between the watershed delineation methods with values from 0.006 to 2%. There are two watersheds that had large percent differences - 08061548 (Figure 3) and 08098450 (Figure 4).

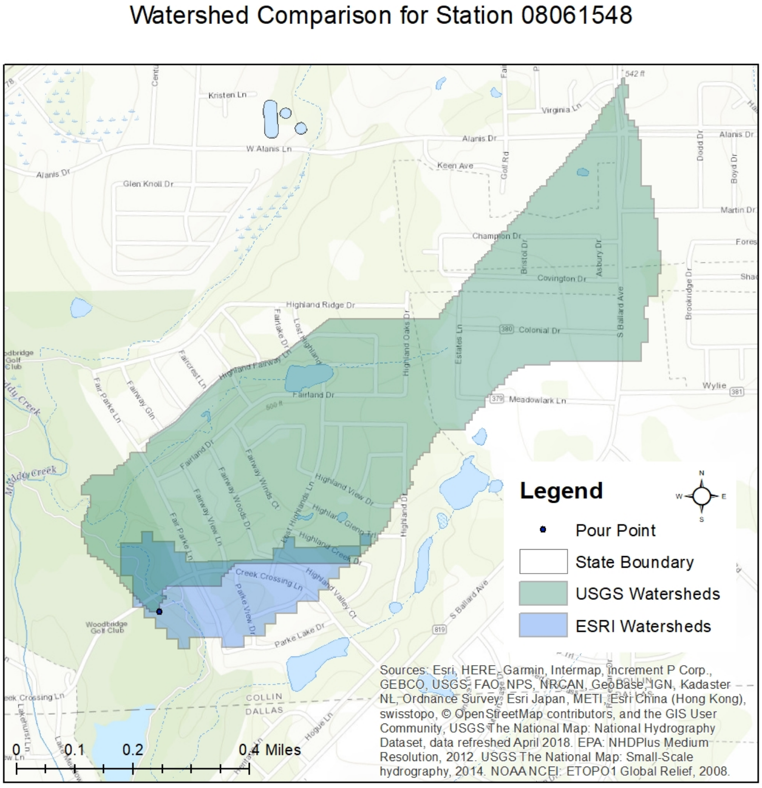


Figure 3: Map of both watersheds for site 08061548.

For watershed 08061548, the percent difference was large at a value of 147%. Because ESRI is a black box software, it is impossible to know exactly what is happening on the back end that causes the two watersheds from the same NWIS site to be so different, but assumptions can be made. First, the spatial resolutions for both watersheds are clearly different, seen by the ESRI watershed appearing coarser than the USGS watershed. There is 10m resolution data available for the continental United States, yet ESRI does not make use of it, which is a major disadvantage to that method. The ESRI Ready-to-Use watershed tool defaults to using 30m DEMs even when the user requests that the watershed be delineated using the best available resolution. The ESRI method watershed output includes metadata information that the DEM is ”NED 30m processed by Esri”, and it is unclear on what year the DEMs were collected, which could be the cause of area discrepancy. Regardless, this example presents itself as a limitation of the ESRI method for use in smaller watersheds. This is the smallest watershed produced in this study, and from this analysis of twelve watersheds, it seems that the ESRI method is less reliable on smaller watersheds. To fully make this claim, more watersheds of similar size would need to be analyzed to be statistically significant.

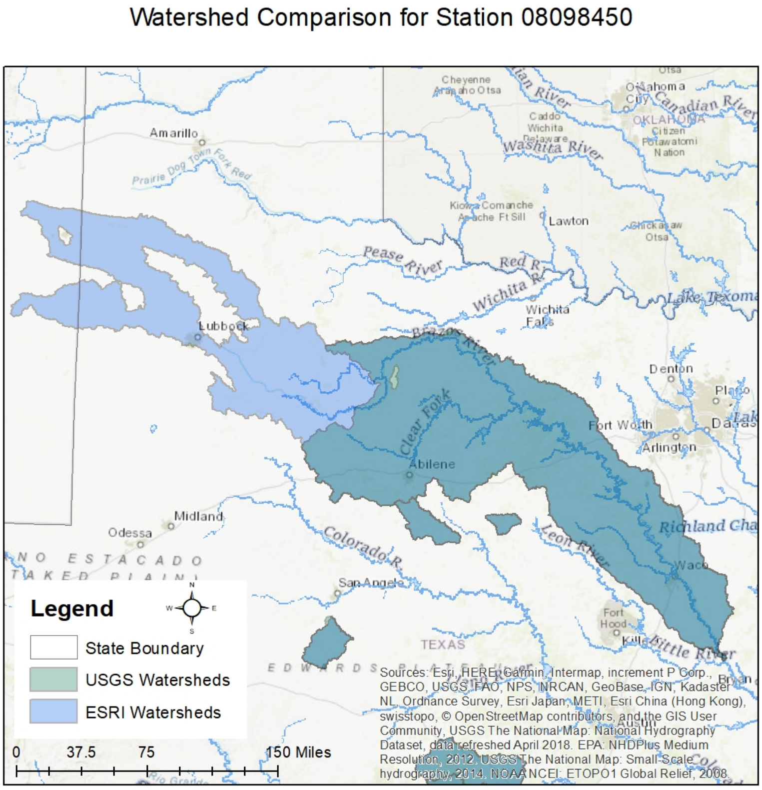


Figure 4: Map of both watersheds for site 08098450.

The other extremely high percent difference presented itself in watershed 08098450 (Figure 4). There is clearly a large difference in basin area of the two watersheds. The ESRI watershed polygon stretches across the Texas-New Mexico border through the Texas panhandle, while the USGS method stops before the Playa Lakes. As stated above, assumptions can be made about why these watershed boundary differences exist. The region in Texas that makes up most of the difference in area between the two methods is called the Playa Lakes and it is a non-contributing drainage area. In the USGS watershed method, the output is a basin area that excludes the non-contributing drainage area. The difference in area between the two methods for this location is caused by the USGS method including only the non-contributing area and the ESRI method including the non-contributing area in addition to the same area as the USGS method (contributing area). This is determined to be a second limitation of the ESRI method. It would be a beneficial and easy-to-implement feature to include an option for the user to choose whether he or she wants the contributing drainage area or non-contributing drainage area, and hopefully this is added as an option in the ESRI tool in the future.

5 Conclusion

Overall, both methods are useful ways for water management to gain information from watershed basins. The USGS method of watershed delineation is the standard that has been used by the USGS for over a decade. It is reliable and produces accurate watersheds, though processing time is not ideal. The watershed outputs from the USGS method are worthy of publication and are trusted by hydrologists to be correct. In contrast to the USGS method, the ESRI Ready-to-Use watershed method has proven to be a useful tool for quick-and-easy basin area visualization, but it is less reliable than the USGS method. This comparison has proven that the ESRI method cannot be implemented by the USGS and its watershed outputs are not worthy of publication. The watershed outputs are in fact limited to being useful tools for visualization to allow the analyst to quickly get an idea of how large or small of a watershed he or she is about to delineate. It is useful in pre-processing of DEMs to be sure all necessary rasters are mosaiced before further processing steps take place. The USGS method will stay as the current method until another potential replacement is identified.

References

[1] National Elevation Dataset, 10 meter resolution.

[2] ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTION ESRI. Esri ArcGIS Pro.

[3] QGIS DEVELOPMENT TEAM (2018). QGIS Geographic Information System.

[4] USGS TEXAS WATER SCIENCE CENTER, GEOSPATIAL SCIENCES AND CYBER INNOVATION. Watershed Delineations Data Release, Texas (*in review*).