AN EVALUATION OF AUTOMATED GIS TOOLS FOR Delineating Karst Sinkholes and Closed Depressions from 1-METER Lidar-Derived Digital Elevation Data

Daniel H. Doctor  
U.S. Geological Survey, 12201 Sunrise Valley Drive, Reston, Virginia, 20192, dhdoctor@usgs.gov

John A. Young  
U.S. Geological Survey, 11649 Leetown Rd., Kearneysville, W. Virginia, 25430, jyoung@usgs.gov

Abstract  
LiDAR (Light Detection and Ranging) surveys of karst terrains provide high-resolution digital elevation models (DEMs) that are particularly useful for mapping sinkholes. In this study, we used automated processing tools within ArcGIS (v. 10.0) operating on a 1.0 m resolution LiDAR DEM in order to delineate sinkholes and closed depressions in the Boyce 7.5 minute quadrangle located in the northern Shenandoah Valley of Virginia. The results derived from the use of the automated tools were then compared with depressions manually delineated by a geologist. Manual delineation of closed depressions was conducted using a combination of 1.0 m DEM hillshade, slopeshade, aerial imagery, and Topographic Position Index (TPI) rasters. The most effective means of visualizing depressions in the GIS was using an overlay of the partially transparent TPI raster atop the slopeshade raster at 1.0 m resolution. Manually identified depressions were subsequently checked using aerial imagery to screen for false positives, and targeted ground-truthing was undertaken in the field. The automated tools that were utilized include the routines in ArcHydro Tools (v. 2.0) for prescreening, evaluating, and selecting sinks and depressions as well as thresholding, grouping, and assessing depressions from the TPI raster. Results showed that the automated delineation of sinks and depressions within the ArcHydro tools was highly dependent upon pre-conditioning of the DEM to produce “hydrologically correct” surface flow routes. Using stream vectors obtained from the National Hydrologic Dataset alone to condition the flow routing was not sufficient to produce a suitable drainage network, and numerous artificial depressions were generated where roads, railways, or other manmade structures acted as flow barriers in the elevation model. Additional conditioning of the DEM with drainage paths across these barriers was required prior to automated delineation of sinks and depressions. In regions where the DEM had been properly conditioned, the tools for automated delineation performed reasonably well as compared to the manually delineated depressions, but generally overestimated the number of depressions thus necessitating manual filtering of the final results. Results from the TPI thresholding analysis were not dependent on DEM pre-conditioning, but the ability to extract meaningful depressions depended on careful assessment of analysis scale and TPI thresholding.

Introduction  
Airborne LiDAR (Light Detection and Ranging) offers enormous potential for mapping sinkholes. However, techniques and approaches for using LiDAR to map sinkholes or other types of depressions have not been standardized. In the past, contour maps or digital elevation models were visually inspected for the presence of closed depressions, and these depressions were identified or manually digitized within a Geographic Information System (GIS) (Angel, 2004; Seale et al. 2008). Visual interpretation and delineation of depression features is painstaking, subjective work. Therefore, datasets of sinkholes and other karst features mapped from GIS data may not be comparable among different interpreters or regions studied. Moreover, complete field verification of individual features is often impractical, thus the reliability of manually digitized sinkhole data produced by even a singular worker may be questionable. Other studies which have examined the use of digital data (including contours derived from LiDAR) for manual interpretation of karst features have shown that subjectivity in the methodology can result in false positive and false negative identification of karst features (e.g., Seale et al., 2008; Vacher et al., 2008).
The goal of this study was to compare manual interpretation to automated detection of sinkholes and other depressions using raster GIS-based methods and a 1m resolution LiDAR-derived ‘bare-earth’ digital elevation dataset. Two approaches of automated depression detection were utilized. The first approach employed GIS reconditioning of the LiDAR DEM for watershed analysis (e.g. ArcHydro tools). Typically, reconditioning is done as a first step to watershed analysis to find and fill areas of closed depressions to their hydrologic spill point in order to correctly model surface flow patterns. Many depressions on coarse-scale DEMs are artifacts of the DEM production process and sink filling removes these errors (Maidment, 2002). However, these depressions are more likely to be actual geomorphic features in areas of karst terrain and on highly accurate LiDAR–derived DEMs (Zandenberg, 2010; Lindsay and Creed, 2006). By subtracting the original elevation data from the resulting filled elevation model, a new difference grid elevation dataset is produced representative of depression location and depths (Anders et al., 2011; Siart et al., 2009; Antonic et al., 2001).

The second approach used the Topographic Position Index (TPI), a GIS moving window operation that calculates the difference between the elevation at each pixel in the DEM and the mean elevation in a neighborhood surrounding the pixel (Jenness et al. 2011). The TPI is similar in concept to other local topographic relief measures that can be calculated in GIS such as the Terrain Shape Index proposed by McNab (1989) and the difference in mean elevation moving window operator proposed by Gallant and Wilson (2000). Different feature scales can be assessed by varying the size of the analysis window, and various feature types can be assessed by using square, triangular, circular, or annular window shapes. The resulting GIS dataset quantifies the landscape position of each pixel as being either higher or lower than a localized average (Jenness et al., 2011). Negative TPI values represent topographic lows (concavities, depressions), while positive TPI values represent topographic highs (convexitities, ridges). The TPI (or similar concepts) has been used recently to find small-scale, concave (e.g., cave) openings (Weishampel et al., 2011), and convex burial mounds (De Reu et al., 2011), and thresholding the TPI values has been suggested as a means of classifying depressional features on the landscape (Klingseisen et al., 2008).

For this study, no effort was made to distinguish among sinkholes and other types of depressions, whether natural or manmade. The goal was simply to test the reliability of automated techniques for finding depressions in a LiDAR elevation model against a manually produced dataset within a GIS. Further work to build a karst feature dataset of the area will focus on separating true karst features from manmade features, and sinkholes from other types of natural depressions such as ponded springs, estavelles, or suffosion depressions.

**Study Area**

The study area is the northern half of the Boyce 7.5 minute quadrangle in Clarke County, Virginia, a region of approximately 70 km$^2$ (Figure 1). The Boyce quadrangle is located within the Shenandoah River drainage basin, an extensive karst region within the Great Valley physiographic province of the Appalachian mountain range. The geology of the quadrangle was originally mapped by Edmundson and Nunan (1973); however, an inventory of karst features was not included in the original mapping. Hubbard (1983) identified a small
number of only the largest depressions in the region from stereophotography and 20-foot contour 1:24,000 scale topographic maps.

The rocks of the region are composed of Paleozoic sedimentary carbonate and siliciclastic rocks that were faulted and folded during the Alleghanian orogeny, and erosion has left the core of the ancient Appalachian mountain range exposed at the present land surface. Karstification in the study area has resulted in a mature, dissected karst surface of moderate to low relief, with 90 m total elevation range and a mean elevation of 180 m above sea level. Sinkholes and karstic depressions generally occur as a result of cover-collapse or suffosion processes within the residuum overlying the carbonate bedrock; the thickness of residuum varies between 0 to upwards of 10 m. Bedrock structures (folds, faults, and fractures) exert strong control over the locations of karstic depressions in the region (Doctor et al., 2008; Doctor and Doctor, 2012).

Methods
Manual delineation of depressions
Airborne LiDAR was acquired over the study area between March 01 and March 09, 2011. Flights took place during leaf-off vegetation conditions and within one week of complete melting of snow cover. Conditions were thus ideal for obtaining a representative bare-earth elevation surface. The vertical accuracy of the delivered LiDAR data was within 9.0 cm root mean square error (RMSE) and the LiDAR point cloud was acquired at a nominal point spacing of 1.0 m; hydro-flattening breaklines were manually collected based on the LiDAR surface model. Although hydro-flattening was performed on areas of known water surfaces, depressions or sinks in the bare-earth elevation model were not filled in the delivered LiDAR data.

The bare-earth LiDAR elevation model was used to derive 1) a hillshade raster, 2) a slopeshade raster, and 3) a TPI raster. The hillshade raster was illuminated from azimuth 300 and an inclination of 20 degrees above the horizon and represented using a grayscale color ramp. The slopeshade raster applied a grayscale color ramp to the surface according to the slope of each pixel. The TPI raster was calculated using a circular annular window with an inner radius of 2 m and outer radius of 10 m (see Jenness et al., 2011 for details), and a grayscale color ramp applied to the resulting raster. These three rasters were overlain upon a 1 m resolution orthorectified aerial imagery of the study region obtained in the years 2002 and 2008 in a GIS using ArcMap® v. 10.0 software. Identification of closed depressions was conducted by close visual inspection within gridded areas defined by the LiDAR tile index; each grid tile region was 1.5 km square.

Depressions in the elevation surface were manually outlined by a single worker (the first author) with a minimum of five vertices per polygon. Upon digitization, each polygon was attributed with the method of recognition (from the field, the LiDAR elevation model, aerial imagery, or some combination of these), and some basic descriptive notes. Description of an individual feature was aided by toggling between LiDAR-derived rasters and georeferenced aerial imagery, and some aspect of the hydrologic condition was normally noted (e.g., ponded, ephemerally ponded, located within an ephemeral channel, takes water, flows as spring, etc.). In general, it was found that the LiDAR rasters were more reliable in terms of an accurately georeferenced dataset than were the aerial images, thus digitization was done on top of the LiDAR-derived imagery. Nonetheless, the georeferenced aerial imagery was critical for identifying false depressions that arose out of processing of the LiDAR bare earth elevation model, as noted by others (e.g., Seale et al., 2008).

Digital outlining of polygons created to represent depressions was best facilitated with the TPI raster at 40% transparency draped over the slopeshade raster (Figure 2). This combination of LiDAR-derived images best illuminated the variations in surface elevation in a manner that could be most readily outlined, avoiding pitfalls of a traditional hillshade surface such as deep shadows, overly highlighted areas, or poor illumination angle and inclination. Although the slopeshade raster helps define edges of features by darkening pixels of high slope (e.g., the edges of a collapsed sinkhole), it does not distinguish between regions of higher or lower elevation; thus, a haystack may appear similar to a sinkhole. Fortunately, the TPI raster permitted distinguishing between areas of higher or lower local elevation (darker areas are local lows, lighter areas are local highs). Thus, the combined use of slopeshade and TPI rasters greatly enhanced the visualization of depressions for manual digitization.
Figure 2. Comparative views of an area with multiple closed depressions in the Boyce quadrangle. In the lower left of each image is a cluster of manmade percolation pits (1); in the upper center are two sinkholes, one deeper than the other (2); in the upper right is a bedrock ridge surrounded by trees along a fence line (3). (A) Aerial image. (B) Hillshade of LiDAR DEM. (C) Slopeshade of LiDAR DEM. (D) TPI raster of LiDAR DEM (note: dark areas are depressions and light areas are ridges). (E) TPI draped over the slopeshade. (F) TPI draped over slopeshade, with manually delineated closed depressions. A red X indicates a false depression where building edges caused errors in the LiDAR elevation model.
Once the polygon data digitization was complete, the vector data were further attributed with geometric characteristics including area, perimeter, major and minor axis length (of an idealized ellipse), circularity index (deviation from area/perimeter ratio of a circle), and elliptical eccentricity. These geometric characteristics were examined to identify threshold values that could be used to filter out spurious depressions identified through the automated techniques.

The eccentricity of an ellipse (commonly denoted as \(e\)) is calculated as:

\[
e = \sqrt{1 - \frac{b^2}{a^2}} \tag{1}
\]

where \(a\) and \(b\) are one-half of the ellipse's major and minor axes, respectively. The eccentricity of an ellipse will be greater than 0 (a perfect circle) and less than 1. Thus, the elliptical eccentricity can be a useful measure of the shape of a possible depression. Based on the manual dataset, we found that an eccentricity of less than or equal to 0.98 worked as a good threshold for identifying true depressions. This eliminated many elongated depressions that appear within stream channels, road ditches, and other features that are unlikely to be true natural depressions.

Similarly, the circularity index is a measure of the deviation of a polygon from a perfect circle based upon its perimeter and area. Since a circle has the smallest perimeter to area ratio, a relationship can be established between the expected circular perimeter of a feature (based on area) and its measured perimeter to create an index of circularity (\(\text{Circ}_c\)):

\[
P_e = 2\pi \left( \frac{\sqrt{A}}{\pi} \right) \tag{2}
\]

\[
\text{Circ}_c = \left( \frac{P_o - P_e}{P_e} \right) + 1 \tag{3}
\]

Where \(A\) is area, \(P_e\) is the expected perimeter if the feature were a perfect circle, and \(P_o\) is the observed perimeter. After some experimentation, we used a circularity threshold of less than 1.7 to capture closed depressions while eliminating linear features.

The smallest area of any of the manually delineated depressions was 7.3 m\(^2\), and the smallest area of the field-checked depressions was 9.0 m\(^2\). In the following automated analyses, we conservatively chose an area cutoff of a true depression to be greater than or equal to 9.0 m\(^2\).

**Automated generation of depressions by difference raster**

Several useful tools for processing elevation data exist in ArcGIS. For most raw elevation data it is necessary to preprocess or recondition the DEM in order to create a ‘hydrologically correct’ elevation model. A hydrologically correct elevation model is one in which every pixel in the surface slopes continually down gradient and out the edges of the elevation model boundaries. If a pixel (or region of pixels) is at a lower elevation than all of the surrounding pixels, the pixel acts as a ‘sink’, and surface flow will stop at that point unless the elevation is raised to a level at which flow would spill out of the sink. Therefore, reconditioning a DEM primarily involves three steps: 1) filling in sinks to their spill level (or ‘pour point’), 2) determining the flow direction within each pixel once the sinks in the DEM are filled, and 3) determining the flow accumulation of each pixel in the elevation model. The flow accumulation raster can then be reclassified to define streams, or those pixels that accumulate the most flow.

The first step in the process uses the Fill tool (under the Spatial Analyst Tools → Hydrology → Fill) in ArcMap v. 10.0. Reconditioning the bare-earth LiDAR DEM with this tool results in a new elevation surface with all sinks filled to their spill elevation. Since a filled pixel may still act as a sink if located within a larger depression, the process is iterative until all pixels within the depression are filled and the depression spills over. As a result of this process, stream channels in the DEM that pass under roads through culverts or beneath bridges may become dammed if the culvert or bridge has not been represented (i.e., cut into the elevation surface to stream level) in the DEM (Figure 3). For most larger streams and rivers, bridges and culverts are normally removed from the DEM during processing of the bare earth elevation model; however, in order to create a hydrologically correct DEM, all of the sinks within it need to be filled, thereby removing any evidence of natural depressions. As requested, the Shenandoah Valley LiDAR was not processed to have sinks filled by the vendor, thus necessitating the process by the end user. This presents an opportunity to partially
automate the process of locating natural depressions. It also represents a challenge for determining true sinks.

Poppenga et al. (2011) suggested a workflow utilizing a least cost path analysis in GIS as an automated means of conditioning a DEM to remove these artificial dams; however, this process does not directly discriminate between true depressions and those that are artifacts of the Fill process, thus all depressions in a DEM will be drained unless selected otherwise according to some criteria. Poppenga et al. (2011) suggested using a threshold area of depressions targeted for drainage greater than or equal to 1,000 square meters, and depth greater than or equal to 1 meter and greater than 0.5 standard deviation of the difference grid. However, these thresholds may include true natural depressions in some karst areas, and exclude a number of depression artifacts. In order to preserve true depressions, the best means is to manually digitize polyline features corresponding to bridges and culverts at known locations, and ‘burn’ these lines into the DEM, or lower the elevation of the DEM along the linear feature to allow flow to pass through the obstructions. This was the approach taken in this study, and was accomplished using the DEM Reconditioning tool in ArcHydro Tools 2.0 (Maidment, 2002). Locating and digitizing actual culverts was facilitated using aerial imagery, and was assisted by the shape of the depression artifacts themselves such that any depression with a flat side parallel to a linear feature such as a road, railway, or driveway was immediately suspect, and targeted to be drained by addition of a culvert.

Once the reconditioned DEM was filled,

1. The original elevation DEM was then subtracted from the filled DEM to generate a ‘fill-difference’ raster that represented the depth of depressions in the original surface.

2. All values in the fill-difference raster that were greater than 0.10 m (10 cm) were extracted to a new raster. The minimum depth threshold of a pixel was conservatively chosen to be 10 cm, or slightly greater than the vertical accuracy of the LiDAR bare earth model which has a RMSE of 9 cm.

3. The values in this threshold fill-difference raster were re-classified to an integer type raster such that pixels in the fill-difference raster with values of less than 0.1 m (10 cm) were classified as ‘No Data.’

4. The remaining pixels were converted into polygons, without simplifying, such that polygon boundaries exactly matched pixel edges.

The result is a polygon layer that is representative of the possible sinks from the reconditioned fill-difference raster based upon the accuracy of the LiDAR data alone. In order to refine these polygons down to those which may represent “true” depressions, we used a training polygon feature class of known depressions that had been identified in the field, and digitally outlined from the LIDAR data. Geometric properties of the true and candidate depression polygons were calculated, and the results were compared.
The geometric properties were calculated as follows:

1. The Zonal Statistics as Table tool was used to extract the MAXIMUM value (depth) from the fill-difference raster overlain by each polygon. The table also contains the area of the polygon, and a count of the number of pixels per polygon (which are both equal values for 1m square pixels).

2. The results of the zonal statistics table were joined to the attributes of the polygons. A threshold value was used to remove polygons with a maximum depth less than 18 cm. This allowed for a great number of possible artificial depressions to be culled from the fill-difference data, and provides a 95% confidence level that a given pixel is a true depression in the elevation model.

3. The Zonal Geometry of each of the polygons was calculated and each polygon was attributed with its area, perimeter, and major and minor axis length of an idealized ellipse that would contain the polygon.

4. The eccentricity of an ellipse was calculated as a measure of elongation of a potential sink, with values between 0 and 1. A threshold value of 0.98 was chosen for the eccentricity based on visual examination of the manually delineated polygons, and any depressions above this threshold were removed.

5. The circularity index was calculated. A threshold value of less than 1.7 was used in order to compare to the elliptical eccentricity results.

**Automated generation of depressions by Topographic Position Index (TPI)**

While the TPI is useful for guiding visual interpretation, it is also useful for quantitative analysis and classification. We examined the potential for mapping closed depressions by thresholding, grouping, and selecting appropriate ranges of TPI values. We calculated TPI across the study area using the Topographic Position Index ArcGIS tool (Jenness et al., 2011) with an annular (i.e. doughnut-shaped) region. We used an inner radius of 2 meters and an outer radius of 10 meters to capture small closed depressions based on previous analysis in Jefferson County, West Virginia (Young, 2007).

Since the TPI calculates the local relative surface elevation of all pixels in the LiDAR DEM (above or below the local mean elevation), a threshold of the TPI values was used such that only depressional elevations are represented. TPI values were summarized within manually mapped, field-checked depression polygons (n=116) to determine appropriate thresholds that would separate significant depressions from shallow swales. TPI thresholds were assessed within size-based classes of field mapped depressions (Table 1). Based on the overall mean of TPI, the TPI threshold was set to distinguish actual depressions as the overall mean TPI value minus one standard deviation, or -0.3 meters.

Additionally, the grouped TPI values characterize linear (e.g. gullies, road cuts) as well as circular depressional features, so we classified and selected features to eliminate non-circular depressions from further consideration. We did this by grouping together pixels from significant depressions into coherent features, and then assessing the shape of these features by computing an index of circularity using area and perimeter relationships in a manner similar to Seale et al. (2008).

By establishing a threshold on circularity of depressional features, elimination of linear or elongated depressions was possible. We also removed from consideration depressions that were located within a 2.0 m buffer zone of building footprints or within major stream channels as false depressions that were likely artifacts of the TPI analysis. Lastly, we set a size threshold of greater than or equal to 9 m² on features to eliminate small, spurious depressions that may have been the result of errors in the bare-earth LiDAR elevation model.

<table>
<thead>
<tr>
<th>Size Class</th>
<th>Mean TPI</th>
<th>SD TPI</th>
<th>Depressions (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 20 m²</td>
<td>-0.1702</td>
<td>0.1627</td>
<td>5</td>
</tr>
<tr>
<td>20-50</td>
<td>-0.1010</td>
<td>0.1385</td>
<td>11</td>
</tr>
<tr>
<td>50-100</td>
<td>-0.1592</td>
<td>0.1206</td>
<td>21</td>
</tr>
<tr>
<td>100-200</td>
<td>-0.1133</td>
<td>0.1698</td>
<td>18</td>
</tr>
<tr>
<td>200-500</td>
<td>-0.1706</td>
<td>0.2243</td>
<td>27</td>
</tr>
<tr>
<td>500-1000</td>
<td>-0.1232</td>
<td>0.1837</td>
<td>14</td>
</tr>
<tr>
<td>100-5000</td>
<td>-0.1034</td>
<td>0.1877</td>
<td>16</td>
</tr>
<tr>
<td>&gt; 5000 m²</td>
<td>-0.0194</td>
<td>0.1576</td>
<td>4</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.1328</td>
<td>0.1740</td>
<td>116</td>
</tr>
</tbody>
</table>

**Table 1.** Mean and standard deviation of topographic position index (TPI) values found at manually mapped, and field checked depressions.
fill-difference method (69%). Of the 36 uncaptured field-checked sinks, all except one were below the 18 cm maximum depth threshold obtained from the fill-difference grid. The one remaining depression had been drained by the addition of a culvert in the reconditioned DEM, and thus ended up being a false negative. Of the other depressions that were too shallow to be captured by the 18cm threshold, 12 additional depressions would have been captured if the threshold were relaxed to within the RMSE of 9 cm, or one standard error.

Overall the semi-automated fill-difference method performed very well in capturing the manually delineated depressions. Before applying area and shape threshold criteria, the total number of depressions identified by the fill-difference method was 3154, nearly four times greater in number than the manually identified depressions. Many of these depressions were unlikely to be true sinks since they were of a very small area. Using 9 m\(^2\) as a conservative area cutoff below which a fill-difference depression would be thrown out, 706 of the fill-difference sinks were removed, and all were less than 60 cm maximum depth. Also, many artificial depressions occur at the edges of structures due to artifacts introduced in the LiDAR processing. Using a polygon layer of building outlines obtained from the Department of Information Technology and GIS for Clarke County, Virginia, the depressions that intersected the building polygons within a 2.0 m buffer distance were removed from the dataset (n=124).

Finally, a large number of depressions appear within stream channels, and although meeting the area, depth, and eccentricity criteria are still artificial depressions within the LiDAR-derived fill-difference dataset. Both the elliptical eccentricity and circularity thresholds are attempts to filter out these artifacts since they account for all but the most linear depressions, which are likely to be spurious. Linear depressions often are observed in gullies, ravines, and stream channels and are not true closed depressions. Using the flow accumulation raster to define stream paths at a threshold flow accumulation of 410,000 (a threshold visually consistent with stream channels in the LiDAR imagery), depressions that directly intersected the stream paths (n=779) were also removed. Applying these threshold criteria of 1) more than 18 cm depth, 2) elliptical eccentricity less than 0.98, 3) area greater than or equal to 9 m\(^2\), 4) not intersecting building outlines, and 5) locations outside of known sinks that were identifiable from the air photo dataset (136), depressions were removed from the remaining dataset (n=746).

Results

In order to evaluate the performance of the semi-automated fill-difference method, the dataset of manually delineated depressions was compared against the fill-difference polygons. Each polygon in the manual dataset was assigned a confidence level as shown in Table 2.

Out of the 842 total manually delineated depressions, 594 were captured by the semi-automated fill-difference method thresholded to 18 cm maximum depth (70%). However, the manually delineated depressions that were greater than 18 cm maximum depth totaled 600. Thus, the semi-automated method captured 99.5% of the depressions it was capable of capturing in the manually created dataset given the 18 cm depth threshold set for removing probable artifacts. In other words, 242 of the manually delineated depressions were shallower than the 18 cm maximum depth threshold used to filter out the fill-difference results. These additional depressions were included in the manual dataset because all depressions that could be seen in aerial imagery were included, regardless of depth. Many of these shallow depressions were likely ponded with water, thus appearing to be shallower in the LiDAR than they truly are due to hydro-flattening or loss of point return intensity.

Of the 6 other manually identified depressions, one (a manmade retention basin) was outside of the 0.98 eccentricity threshold, three were drained during creation of the fill grid, and two were below the 18cm depth threshold in the fill-difference grid. A very small number of depressions (3) were artificially drained during the reconditioning of the fill grid, resulting in false negatives.

Out of the 116 possible field-checked sinks of highest confidence, 80 were captured by the semi-automated fill-difference method (69%). Of the 36 uncaptured field-checked sinks, all except one were below the 18 cm maximum depth threshold obtained from the fill-difference grid. The one remaining depression had been drained by the addition of a culvert in the reconditioned DEM, and thus ended up being a false negative. Of the other depressions that were too shallow to be captured by the 18cm threshold, 12 additional depressions would have been captured if the threshold were relaxed to within the RMSE of 9 cm, or one standard error.

Overall the semi-automated fill-difference method performed very well in capturing the manually delineated depressions. Before applying area and shape threshold criteria, the total number of depressions identified by the fill-difference method was 3154, nearly four times greater in number than the manually identified depressions. Many of these depressions were unlikely to be true sinks since they were of a very small area. Using 9 m\(^2\) as a conservative area cutoff below which a fill-difference depression would be thrown out, 706 of the fill-difference sinks were removed, and all were less than 60 cm maximum depth. Also, many artificial depressions occur at the edges of structures due to artifacts introduced in the LiDAR processing. Using a polygon layer of building outlines obtained from the Department of Information Technology and GIS for Clarke County, Virginia, the depressions that intersected the building polygons within a 2.0 m buffer distance were removed from the dataset (n=124).

Finally, a large number of depressions appear within stream channels, and although meeting the area, depth, and eccentricity criteria are still artificial depressions within the LiDAR-derived fill-difference dataset. Both the elliptical eccentricity and circularity thresholds are attempts to filter out these artifacts since they account for all but the most linear depressions, which are likely to be spurious. Linear depressions often are observed in gullies, ravines, and stream channels and are not true closed depressions. Using the flow accumulation raster to define stream paths at a threshold flow accumulation of 410,000 (a threshold visually consistent with stream channels in the LiDAR imagery), depressions that directly intersected the stream paths (n=779) were also removed. Applying these threshold criteria of 1) more than 18 cm depth, 2) elliptical eccentricity less than 0.98, 3) area greater than or equal to 9 m\(^2\), 4) not intersecting building outlines, and 5) locations outside of known sinks that were identifiable from the air photo dataset (136), depressions were removed from the remaining dataset (n=746).

Table 2. Identification method and confidence of manually identified depressions.

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Identification Method</th>
<th>Depressions (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>air photo</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>LIDAR</td>
<td>453</td>
</tr>
<tr>
<td>3</td>
<td>air photo, LIDAR</td>
<td>229</td>
</tr>
<tr>
<td>4</td>
<td>LIDAR, air photo</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>in field</td>
<td>116</td>
</tr>
<tr>
<td>Total Depressions:</td>
<td></td>
<td>842</td>
</tr>
</tbody>
</table>
stream channels, the number of depressions resulting from the semi-automated fill-difference method was 1695, or roughly twice the number identified through the manual methods (n=842).

By contrast, the thresholding, grouping, and shape identification of the TPI raster resulted in identification of 3980 total depressions; however, only 257 of these coincided with the 842 manually mapped depressions (Table 3). Likewise, only 39 of 116 field mapped depressions were identified using the TPI threshold of -0.3 meters difference from average elevation in the 2-10 meter annular window, a shape threshold (circularity index) of less than 1.7, and a size greater than 9 m². All but 32 of the 3980 TPI depressions thresholded to circularity of 1.7 met the eccentricity criteria of 0.98. In general, the TPI method resulted in more numerous and smaller depressions than either the manual or the fill difference method. We acknowledge that there is redundancy between the elliptical eccentricity and the circularity threshold measures. While the eccentricity measure combined with the fill-difference method appeared to map more “true” sinks than the TPI method and circularity threshold, a more detailed comparative analysis is needed to fully evaluate the strengths and weaknesses of these measures.

### Conclusions

Overall, implementation of the TPI method is much more time efficient than the fill-difference methods, and does not require re-conditioning of the original LiDAR DEM with known culvert locations. However, the fill-difference method was much better at capturing depressions of larger area, and suffered less than the TPI method from identifying artifacts as real depressions in the LiDAR DEM. More work is needed to assess mapping performance at multiple TPI window sizes as the single scale analysis presented here may not capture the variation in depression feature area and shape.

It was found that a number of fill-difference depressions were evident in the results that were likely artificial, but had not been drained due to the uncertainty of culvert locations. Thus, additional work is needed to cull out possible artificial fill-difference depressions ‘dammed’ against roads where culverts were not identified in aerial imagery. Fieldwork targeted at defining culvert locations would likely prove fruitful for mapping sinkholes from LiDAR, perhaps as much as the effort expended to map the locations of depressions themselves.

### References


### Table 3. Results from applying thresholds to the TPI raster for TPI value, shape, and size to identify depressions.

<table>
<thead>
<tr>
<th>Manual mapping method</th>
<th>TPI mapped</th>
<th>Average Area (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPI sink, no manual equivalent</td>
<td>3723</td>
<td>19.89</td>
</tr>
<tr>
<td>1 - air photo</td>
<td>12</td>
<td>22.92</td>
</tr>
<tr>
<td>2 - LiDAR</td>
<td>139</td>
<td>29.88</td>
</tr>
<tr>
<td>3 - air photo, LiDAR</td>
<td>57</td>
<td>34.42</td>
</tr>
<tr>
<td>4 - LiDAR, air photo</td>
<td>10</td>
<td>32.30</td>
</tr>
<tr>
<td>5 - in field</td>
<td>39</td>
<td>27.18</td>
</tr>
</tbody>
</table>

Total: 3980