

Automated delineation of karst sinkholes from LiDAR-derived digital elevation models



Qiusheng Wu^{a,*}, Chengbin Deng^a, Zuoqi Chen^b

^a Department of Geography, Binghamton University, State University of New York, Binghamton, NY 13902, United States

^b Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai 200241, China

ARTICLE INFO

Article history:

Received 12 June 2015

Received in revised form 4 May 2016

Accepted 5 May 2016

Available online 07 May 2016

Keywords:

Sinkhole

Karst

Closed depressions

LiDAR

Contour tree

Minnesota

ABSTRACT

Sinkhole mapping is critical for understanding hydrological processes and mitigating geological hazards in karst landscapes. Current methods for identifying sinkholes are primarily based on visual interpretation of low-resolution topographic maps and aerial photographs with subsequent field verification, which is labor-intensive and time-consuming. The increasing availability of high-resolution LiDAR-derived digital elevation data allows for an entirely new level of detailed delineation and analyses of small-scale geomorphologic features and landscape structures at fine scales. In this paper, we present a localized contour tree method for automated extraction of sinkholes in karst landscapes. One significant advantage of our automated approach for sinkhole extraction is that it may reduce inconsistencies and alleviate repeatability concerns associated with visual interpretation methods. In addition, the proposed method has contributed to improving the sinkhole inventory in several ways: (1) detection of non-inventoried sinkholes; (2) identification of previously inventoried sinkholes that have been filled; (3) delineation of sinkhole boundaries; and (4) characterization of sinkhole morphometric properties. We applied the method to Fillmore County in southeastern Minnesota, USA, and identified three times as many sinkholes as the existing database for the same area. The results suggest that previous visual interpretation method might significantly underestimate the number of potential sinkholes in the region. Our method holds great potential for creating and updating sinkhole inventory databases at a regional scale in a timely manner.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Sinkholes are closed depressions in the Earth's surface with internal drainage caused by subsurface dissolution of soluble bedrock in karst landscapes (Miao et al., 2013). Sudden sinkhole collapse and gradual ground subsidence phenomenon may cause severe damage to human properties and affect water quality in underlying carbonate acquirers (Shaban and Darwich, 2011; Rahimi and Alexander, 2013; Galve et al., 2015; Taheri et al., 2015). Consequently, sinkhole inventory mapping is critical for understanding hydrological processes and mitigating geological hazards in karst areas. The reliability of sinkhole susceptibility and hazard maps and the effectiveness of mitigation activities largely rely on the representativeness, completeness, and accuracy of the sinkhole inventories on which they are based (Al-Kouri et al., 2013; Gutiérrez et al., 2014). In the last few decades, a number of institutions and associations in several states of the United States have developed sinkhole or other karst feature databases mostly integrated in Geographical Information Systems (GIS), including Kentucky (Dinger et al., 2007; Shukunobe, 2012; Zhu et al., 2014), Minnesota (Gao et al., 2002; Larson, 2009; Rahimi and Alexander, 2013), Missouri (Mukherjee,

2012; Nwokebuihe et al., 2014), and Florida (Montane, 2001; Seale et al., 2008; Vacher et al., 2008).

However, most previous methods for mapping sinkholes were primarily based on visual interpretation of low-resolution topographic maps (e.g. U.S. Geological Survey 1:24,000 scale topographic maps) and aerial photographs with subsequent field verification, which are labor-intensive and time-consuming. Moreover, complete field verification of each sinkhole is often impractical, thus the reliability of manually digitized sinkhole data by even the same worker may be questionable (Doctor and Young, 2013). Last but not least, some previous studies (Witthuhn and Alexander, 1995; Rahimi and Alexander, 2013; Zhu et al., 2014) found that sinkholes might be changing fast due to natural or anthropogenic causes such as urban development and agricultural expansion. Therefore, we need to automate mapping of sinkholes to update the sinkhole inventory regularly. In recent decades, the advent of airborne Light Detection and Ranging (LiDAR) and Interferometric Synthetic Aperture Radar (InSAR) remote sensing technologies have produced large volumes of highly accurate and densely sampled topographical measurements. The increasing availability of high-resolution digital elevation data derived from LiDAR and InSAR technologies allows for an entirely new level of detailed delineation and analyses of small-scale geomorphologic features and landscape structures at fine scales (Berardino et al., 2002; Lindsay and Creed, 2006; Gutiérrez et al., 2011; Huang et al., 2014; Wu et al., 2014; Galve et al., 2015).

* Corresponding author.

E-mail address: wqs@binghamton.edu (Q. Wu).

Several studies that employed LiDAR data to identify and characterize sinkholes and other karst features have been reported in the literature. For example, Mukherjee (2012) applied a sink-filling method on LiDAR data to create a depression-less digital elevation model (DEM) in Nixa, Missouri, USA. The original DEM was subtracted from the depression-less/filled DEM to locate the depressions. The subtracted layer was then processed with different thresholds of elevation differences to locate sinkholes. Zhu et al. (2014) used a similar sink-filling method of processing LiDAR data to map sinkholes in portions of the Floyds Fork watershed in central Kentucky, USA, and they found four times as many potential sinkholes as the existing database for the same area. Other methods for karst sinkhole recognition include image filtering techniques with kernel windows using focal functions (Obu and Podobnikar, 2013) and the “active-contour” approach (Rahimi and Alexander, 2013), an algorithm that delineates sinkhole boundaries based on changes in flow of the elevation gradient in the surrounding region around the local minima at potential sinkhole locations.

In this paper, we present an improved method for mapping karst sinkholes by using a localized contour tree method on digital elevation data. The conceptual model of localized contour tree (Wu et al., 2015) based on graph theory (Heckmann et al., 2015; Phillips et al., 2015) was used to identify and delineate surface depressions across different spatial scales and characterize their morphometric properties and topological relationships. The identified depressions were further refined using morphometric parameters to extract potential karst sinkholes. Our application example in Fillmore County, Minnesota, USA, demonstrates that the proposed method is effective and it holds great potential for creating and updating sinkhole inventory databases at a regional scale in a timely manner.

2. Study area and data

2.1. Study area

Our study area, Fillmore County (Fig. 1), is an active karst area (Rahimi and Alexander, 2013) located in southeastern Minnesota, USA, which is part of the Upper Mississippi Valley Karst (Hedges and Alexander, 1985). The county has a total area of 2230 km², which is primarily composed of cultivated crops (41.5%), pasture/hay (21.8%), deciduous forest (18.6%), grassland/herbaceous (11.5%), and developed land (5.4%) according to the National Land Cover Database 2011 (Jin et al., 2013). Most surficial karst features, such as sinkholes, are only found in those areas with less than 15 m of sedimentary cover over bedrock surface

(Gao et al., 2002). Since the 1990s, various efforts have been made by the Minnesota Geological Survey and the Minnesota Department of Natural Resources to map karst features and publish various versions of karst feature distribution maps for southeastern Minnesota (Hedges and Alexander, 1985; Alexander and Maki, 1988; Magdalene and Alexander, 1995; Shade et al., 2001; Gao et al., 2002; Gao and Alexander, 2003; Gao et al., 2005; Larson, 2009; Rahimi and Alexander, 2013). Gao et al. (2002) initiated the development of the Minnesota Karst Feature Database (KFDB) for southeastern Minnesota that allows sinkholes and other karst features (springs, stream sinks, etc.) to be displayed and analyzed in a GIS environment. The KFDB is provided as point features in ESRI Shapefile format (Minnesota Geospatial Commons, 2005).

According to the metadata, the latest update to the KFDB was conducted in November 2005, a decade ago. As of November 2005, 9128 sinkholes have been mapped and recorded in southeastern Minnesota. Of the 9128 sinkholes in the KFDB, 6139 (67.3%) were located in Fillmore County. Fig. 1 shows the sinkhole distribution overlain on the LiDAR DEM shaded relief in Fillmore County. It appears that the majority of sinkholes were highly concentrated on the flat hilltops between or adjacent to river valleys. The flat hilltops in a northwest to southeast band across the central part of Fillmore County are part of an old erosion surface that cuts across the stratigraphy. In Fillmore County, limestone and dolostone underlie most of the County. Fillmore County has been called the “Karst Capital of Minnesota” (Witthuhn and Alexander, 1995). With 6139 sinkholes and 873 springs that have been mapped, the county is believed to have more karst features than the rest of southeastern Minnesota combined. However, Witthuhn and Alexander (1995) found that sinkholes are forming rapidly in southeastern Minnesota due to both natural and anthropogenic processes. They estimated that the rate of sinkhole formation was about 2% per year of the total inventory of sinkholes. This high rate of formation indicates that many sinkholes are ephemeral features in the landscape, which calls for the needs for regular updating of sinkhole database in the region. A recent study by Rahimi and Alexander (2013) revealed that approximately two-thirds of the inventoried sinkholes in Winona County in southeastern Minnesota have been filled for agricultural use and other reasons.

2.2. LiDAR data

The LiDAR data for Fillmore County were acquired during November 18–24, 2008 as part of the LiDAR data acquisition project for nine counties in southeastern Minnesota (Minnesota Geospatial Commons, 2008). The bare-earth LiDAR-derived DEM is in the map projection of

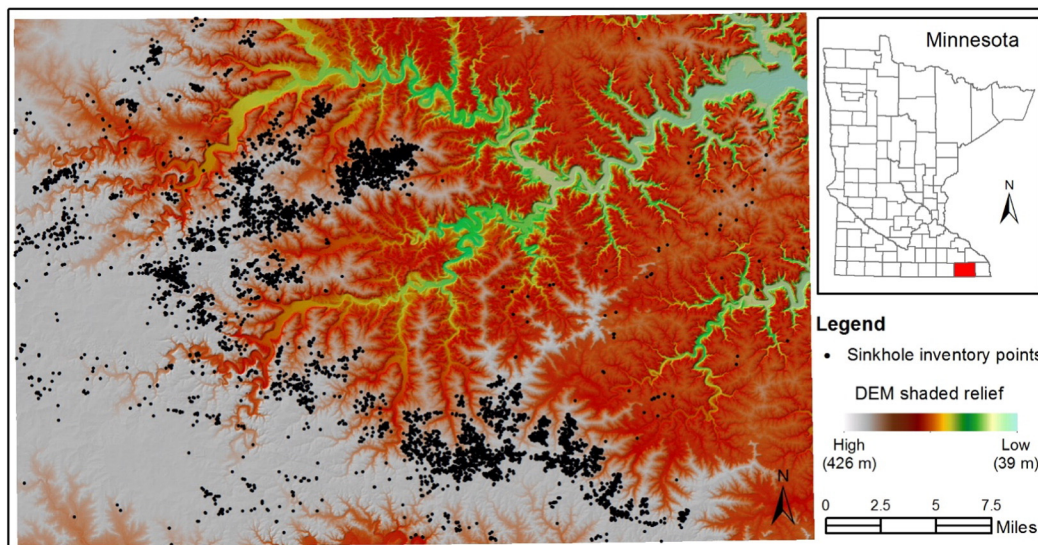


Fig. 1. Distribution of sinkhole inventory points in Fillmore County, Minnesota, USA.

Universal Transverse Mercator (UTM) Zone 15N and referenced to horizontal datum-NAD83 and vertical datum NAVD88. The LiDAR DEM has 1 m spatial resolution and its *RMSE* (root mean squared error) of vertical measurements was estimated to be 28.7 cm at a 95% confidence level of all land cover categories. The shaded relief map of the bare-earth LiDAR DEM is shown in Fig. 1. The elevation values range from 39 to 426 m, with an average elevation of 362 m. In addition to the LiDAR data, we used leaf-off 4-band (red, green, blue, and infrared) digital orthoimagery for validating our sinkhole detection results. The digital orthoimagery for Fillmore County was flown in mid-April 2011 at 0.5 m resolution. Both data sets can be obtained from the Minnesota Geospatial Commons website (Minnesota Geospatial Commons, 2015).

3. Methods

3.1. Outline

Our methodology for karst sinkhole mapping is a semi-automated approach involving several steps: (a) LiDAR DEM preprocessing; (b) depression identification using the localized contour tree method; (c) calculation of morphometric properties of depressions; and (d) sinkhole extraction by eliminating non-sinkhole depressions using morphometric parameters. The flowchart in Fig. 2 shows the detailed data processing steps.

3.2. LiDAR DEM preprocessing

In our study, the basic assumption for sinkhole detection is that sinkholes are a subset of surface depressions, which might also include other non-sinkhole natural depressions or man-made depressions. Since the localized contour tree method is a vector-based method applied to the LiDAR DEM, the purpose of this preprocessing step is to extract a subset of the original LiDAR DEM that represents surface depressions. In this way, those non-depression areas will be eliminated from the analysis as they are unlikely to be sinkholes. This can greatly reduce the number of contours being generated and thus reduce the computation time.

As data noise or errors in the DEM may lead to jagged, irregular or fragmented contour lines (Wu et al., 2015), a 3×3 median morphological operator was used to smooth the LiDAR DEM. The median operator is an edge-preserving filter that is used to remove data noise and suppress small artifact depressions without distorting the boundaries of true surface depressions, and is considered better than a mean (averaging) filter (Liu et al., 2010; Wu et al., 2014). After smoothing the original DEM with the median filter, we used the efficient depression filling algorithm developed by Wang and Liu (2006) to generate a new filled DEM. The algorithm identifies and fills surface depressions by

introducing the concept of spill elevation and integrating the priority queue data structure into the least-cost search of spill paths. It has been widely adopted and implemented in several GIS software packages due to its high computation efficiency and coding simplicity (Lindsay, 2014). By subtracting the original DEM from the resulting filled DEM, a new elevation difference grid is generated representative of depression location and depths. The elevation difference grid was converted into polygons without boundary simplifications in order to make the polygon boundaries exactly match pixel edges. After converting raster to polygon, the polygon should undergo a buffer analysis to make sure most of the contour lines are closed, especially at the edge of the study area. The polygon layer is used as a mask to extract a subset of the smoothed LiDAR DEM representing depression regions.

3.3. Depression identification

Based on the subset of the smoothed LiDAR DEM representing depression areas, we generated vector contours by setting the base contour to be 39 m and the contour interval to be 0.5 m, which is slightly greater than the vertical accuracy of the LiDAR DEM with an *RMSE* value of 28.7 cm. The localized contour tree method (Wu et al., 2015) was then applied to the contours to identify surface depressions. We used a minimum depression area of 100 m² and minimum depression depth of 0.5 m, which allows identification of sinkholes that are larger than 100 m² and deeper than 0.5 m. We considered this chosen area and depth thresholds sufficient to identify most natural sinkholes in the study area.

In the contour maps, depressions are represented as closed contours that are surrounded by other closed contours at a higher elevation. Consequently, only closed contours are kept for further analysis while open contours that do not form a loop are eliminated from further analysis. Topology between closed contours is then constructed. Specifically, each closed contour is attributed with its adjacent outward contour, if any, and the corresponding contour elevation. To facilitate the algorithm for fast searching of depressions, one key step is to identify the so-called “seed contours”, which are defined as closed contours that do not enclose any other contours (Wu et al., 2015). Fig. 3a shows an example of a compound surface depression, and the elevation profile of the transect A–B is shown in Fig. 3b. It appears that the compound depression has two smaller depressions nested inside. As shown in Fig. 3a, contours M and O are seed contours whereas contours N, P, Q, and R are not. The seed contours serve as the starting point to search outwards for other associated closed depression contours, which are identified as the 1st rank contours. Assuming that the water level increases inside the depression, the water will spill at a certain threshold, which is named as the spilling elevation of a depression (Wang and Liu, 2006). In other words, the spilling elevation of a depression has the highest elevation for water ponding. Beyond the spilling elevation, the ponded water spills. If two or more adjacent depressions (1st rank) share the same spilling elevation, they will be merged and form a 2nd rank depression (see Fig. 3a). Similarly, a further combination of 2nd rank depressions forms even higher level depressions. Whenever two or more depressions merge, a higher rank of depression is created.

The topological relationships between depressions (contours and their bounded regions) can be represented by a contour tree. The concept of contour tree was first proposed by Kweon and Kanade (1994) for representing the spatial relations of contour lines where contours are mapped as nodes and interstitial spaces as links. The nodes in the tree graph represent contours, and the link (edge) between nodes represents the adjacency and containment relationships between contours. In the contour tree, splitting and merging of nodes represent the change in topology. As shown in the contour tree graph in Fig. 3a, contour (depression) N encloses the depression seed contour M (1st rank), and there is no topological change between them. Therefore, contour N is also determined as the 1st rank contour, the same rank as the depression seed contour M. With the water level increasing, the depressions (contours) N and O would merge to form a 2nd rank depression P.

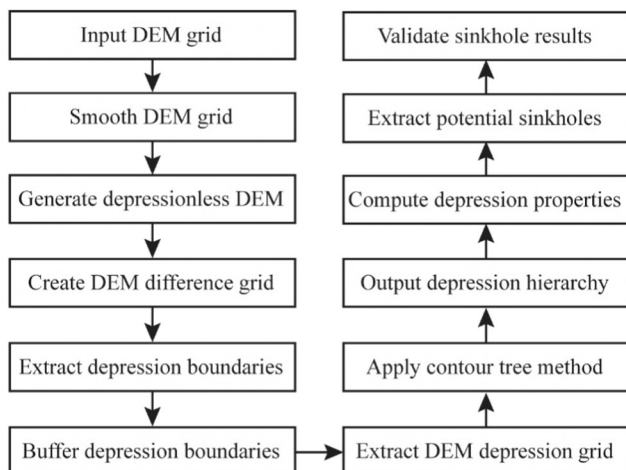


Fig. 2. Flowchart of the methodology.

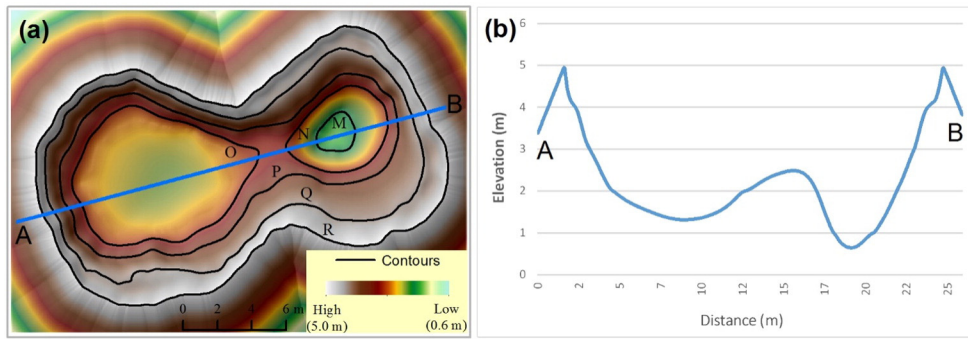


Fig. 3. Contour representation of a compound surface depression. (a) Contours overlain on DEM shaded relief. (b) Elevation profile of the transect A–B shown in (a).

This iterative procedure continues until all the depression seed contours and their outward closed contours are processed and depression ranks are determined accordingly.

In the perspective of graph theory, the hierarchical relationships of nested depressions inside a compound depression constitute a tree. The most outward contour of the compound depression is the root of the tree, the directed links between two adjacent contours are the edges of the tree, and seed contours are the leaf nodes of the tree. The depression seed contours are used as the starting point to search outwards to minimize search time for establishing the tree or forest of trees. Compared to the global contour tree method (Wu, 2011), the localized contour tree method is more effective and computationally efficient. Instead of creating a single global tree for the entire area, the localized contour tree algorithm constructs a forest of trees. Each tree represents one compound depression, and the number of trees in the forest represents the number of compound depressions for the entire area. A simple surface depression (1st rank) constitutes a single-branch contour tree, while a compound surface depression is represented by a multi-branch contour tree. For example, the corresponding contour tree for the compound depression shown in Fig. 3a is a multi-branch tree (Fig. 4a). To explicitly represent the nested hierarchy of a compound surface depression, the contour tree is simplified by removing those nodes without topological change. Only those nodes with topological changes (immediately before merging) are kept in the simplification. It should be noted that the root node of the contour tree is always kept in the simplification since it represents the maximum boundary of the compound depression. After simplification, the multi-branch contour tree shown in Fig. 4a is reduced to a smaller compact tree with only three nodes (Fig. 4b). The leaf nodes of the simplified contour tree represent two simple depressions (N and O) at the 1st

rank. The parent (root) node represents the compound depression (R) at the 2nd rank. The simplified contour tree gives a clear representation of the nested hierarchical structure of a compound surface depression.

3.4. Calculation of depression characteristics

After identifying surface depressions and quantifying their ranks according to their topological relationships, we calculated the basic morphometric characteristics for each depression at each rank, including the width (w), length (l), area (A), perimeter (p), depth, volume, elongatedness (ELG), compactness index (CI), and standard deviation of elevation (STD). The method by Chaudhuri and Samal (2007) was adopted to compute the minimum bounding rectangle for each depression polygon. The depression length was defined as the length of the major axis and the depression width was defined as the length of the minor axis of the fitted minimum bounding rectangle (Fig. 5). The perimeter is the length of the contour that delimits the sinkhole in plain view, and the sinkhole area is considered as the planimetric surface bounded by the perimeter. Depth is defined as the maximum depth between the sinkhole edge and the deepest point within the sinkhole. STD calculates the standard deviation of all cells in the DEM that belong to the same depression.

ELG is defined as the ratio between the length and width of the fitted minimum bounding rectangle:

$$ELG = \frac{l}{w}. \quad (1)$$

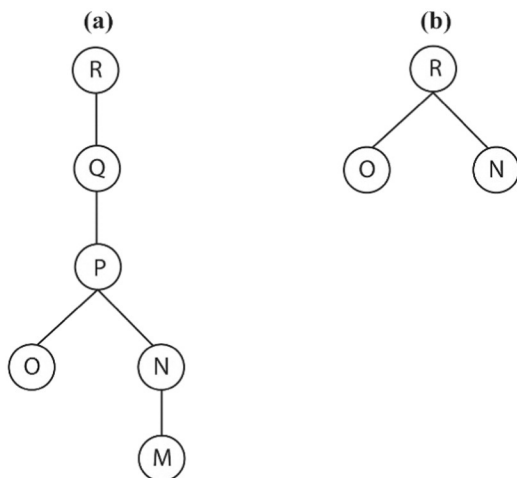


Fig. 4. Contour tree representation of the compound surface depression shown in Fig. 3. (a) Regular contour tree representation. (b) Simplified contour tree representation.

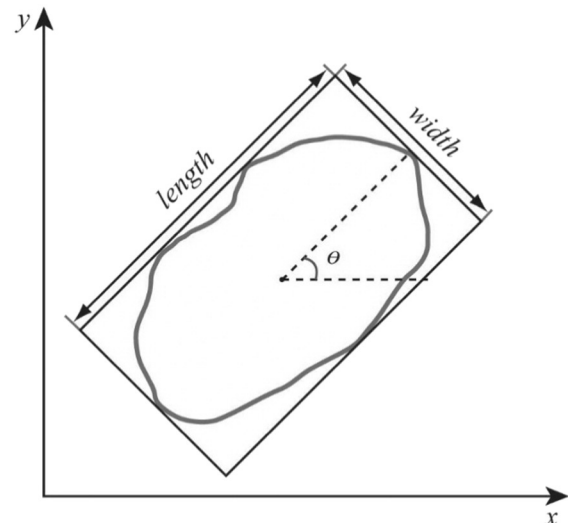


Fig. 5. Fitted minimum bounding rectangle for a surface depression.

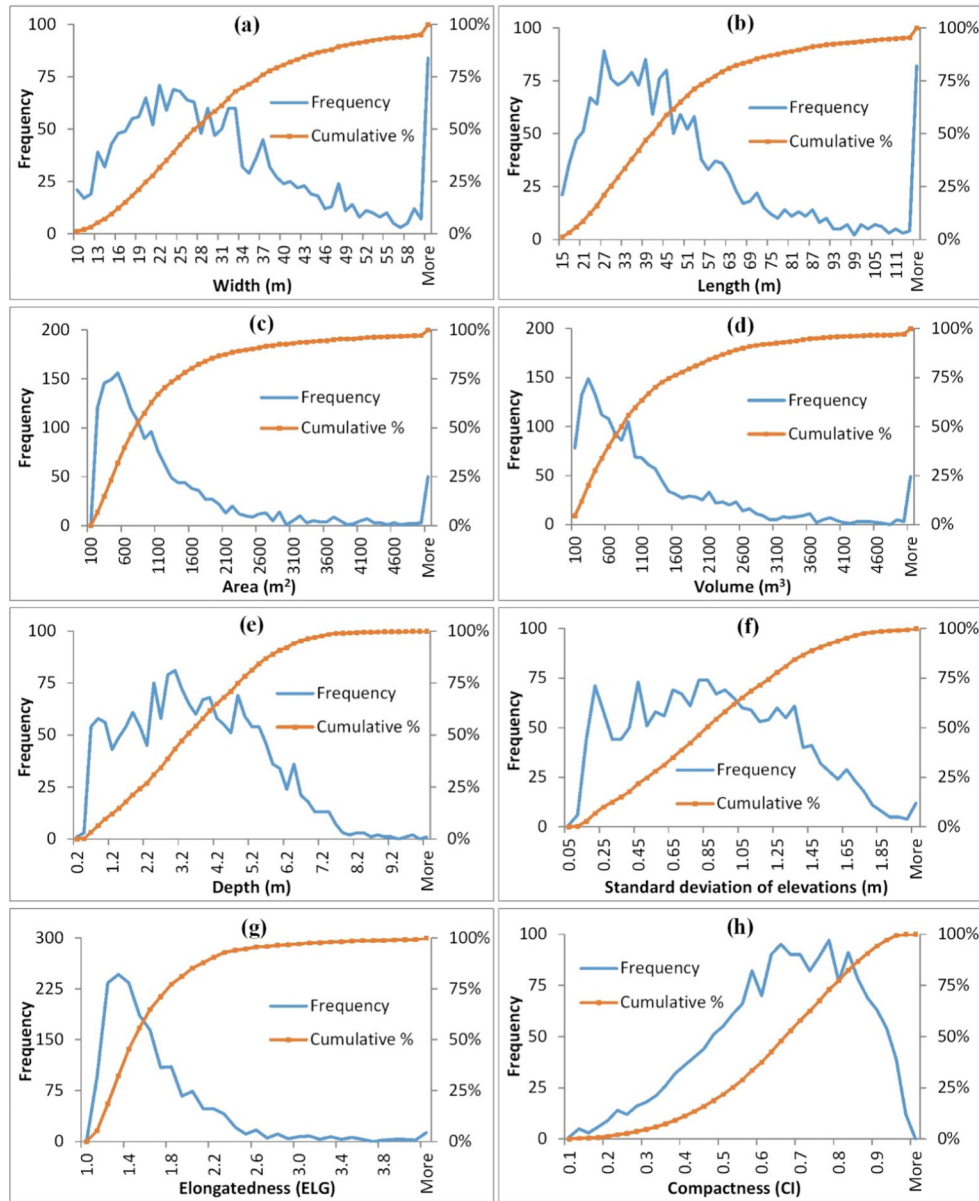


Fig. 6. Histograms of morphometric characteristics of 1784 training sinkholes (1st rank). (a) Width. (b) Length. (c) Area. (d) Volume. (e) Depth. (f) Standard deviation of elevations. (g) Elongatedness. (f) Compactness.

A circle and square will have the smallest value for *ELG*. [Basso et al. \(2013\)](#) classified sinkholes into four groups according to *ELG*: (i) circular and sub-circular ($ELG \leq 1.21$); (ii) elliptical ($1.21 < ELG \leq 1.65$); (iii) sub-elliptical ($1.65 < ELG \leq 1.8$); and (iv) elongated ($ELG > 1.8$).

Table 1
Summary statistics of 1784 training sinkholes (1st rank).

Statistics	Width (m)	Length (m)	Area (m ²)	Volume (m ³)	Depth (m)	STD (m)	ELG	CI
Min	7.4	12.8	101.0	15.6	0.17	0.05	1.01	0.06
5th percentile	12.9	18.6	169.2	108.3	0.73	0.18	1.10	0.31
Median	27.1	40.7	751.0	800.4	3.37	0.84	1.44	0.66
Mean	30.8	49.2	1225.6	1271.5	3.48	0.87	1.60	0.64
95th percentile	59.0	110.8	3696.9	3667.1	6.55	1.64	2.54	0.91
Max	236.1	381.7	30,029.0	56,128.5	13.13	3.72	11.20	0.97

CI is a widely used shape indicator ([Davis, 2002](#)) defined by the perimeter and area of the object:

$$CI = \frac{4\pi A}{p^2}. \quad (2)$$

The most compact object in a Euclidean space is a circle. A circle-shaped object has a compactness index of unity. The compactness index is also known as the circularity measure ([Pratt, 1991](#)).

The abovementioned morphometric properties were examined to identify threshold values that could be used to filter out non-sinkhole/spurious depressions.

3.5. Sinkhole Extraction Analyst

To streamline the procedures for automated sinkhole extraction, we have implemented the proposed method as an ArcGIS toolbox – Sinkhole Extraction Analyst, which will be freely available for the public to

download free of charge in the near future. The core algorithm components are developed using the Python programming language. The toolbox includes two tools: Depression Identification Tool and Sinkhole Extraction Tool. The Depression Identification Tool asks the user to provide a single input, the LiDAR DEM, and then executes the aforementioned procedures with user-specified parameters such as the base contour, contour interval, minimum depression area, and minimum depression depth to automatically create depression polygons at different ranks. The Sinkhole Extraction Tool selects and exports potential sinkholes based on user-specified criteria related to the area, depth, *STD*, *ELG*, *CI*, etc. All depression and sinkhole results can be saved as ESRI Shapefile or Geodatabase format.

4. Results

Using the localized contour tree method with certain parameter values (contour interval = 0.5 m; base contour = 39 m; and minimum area = 100 m²), we identified 14,499 depressions at 1st rank from the LiDAR-derived DEM in Fillmore County, nearly three times greater in number than the inventoried sinkholes in the KFDB. The number of 2nd, 3rd, 4th, and 5th rank depressions were 1668, 235, 75, and 17, respectively. Depressions detected using the localized contour tree is a combination of sinkholes and other natural or man-made depressions, including stream channels, ponds, retention basins, and road ditches. The next step is to extract potential sinkholes from these detected depressions.

Sinkholes vary in size, shape and distribution in different regions of the world. Consequently, sinkholes are defined differently in literature. For instance, Kobal et al. (2014) described karst solution sinkholes in southwest Slovenia as having more than 2 m deep basins with more than 10 m diameter. Mukherjee (2012) used a 4-m depth threshold to locate sinkholes in Nixa, Missouri, while Zhu et al. (2014) considered a 6-m depth threshold sufficient to identify most natural sinkholes in the Floyds Fork watershed in central Kentucky, USA. In Fillmore County, Witthuhn and Alexander (1995) reported that sinkholes ranged from <1 m to >30 m in diameter and 0.3–18 m in depth; the majority of them were 3–12 m in diameter and 1.5–12 m deep. This estimation was based on a limited number of sinkholes surveyed in the field as only sinkhole locations have been recorded in the KFDB. No data about sinkhole boundaries for Fillmore County has been reported in

the literature. By intersecting the KFDB sinkhole points layer with the depression polygons layer generated from the localized contour tree method, we found that 1858 (30.3%) out of 6139 inventoried sinkholes in the KFDB were located within 1784 depressions (1st rank). These intersected depressions were considered as the “training sinkholes” in our study, as we believed they were very likely to be true sinkholes due to coincident locations with the KFDB. The histograms and summary statistics of morphometric properties of the 1784 training sinkholes are shown in Fig. 6 and Table 1, respectively.

Based on summary statistics of the 1784 training sinkholes, it appears that most sinkholes ranged from 169 m² (5th percentile) to 3696 m² (95th percentile) in size and from 0.73 m (5th percentile) to 6.55 m (95th percentile) in depth. The median size and depth are 751 m² and 3.37 m, respectively. The former is approximately equal to the size of a circle with 30 m in diameter, which can be evidenced from the median width (27.1 m) and length (40.7 m) based on the minimum bounding rectangle. The standard deviation of elevations within each depression ranged from 0.18 m (5th percentile) to 1.64 m (95th percentile). Most training sinkholes had *CI* values greater than 0.31 and *ELG* values less than 2.54.

In order to refine the detected depressions down to those which may represent “true” sinkholes, we used a combination of morphometric parameters based on the summary statistics of the 1784 training sinkholes: area < 4000 m², depth > 0.5 m, *STD* > 0.18 m, *ELG* < 2.54, and *CI* > 0.31. Since most natural sinkholes tend to have circular or elliptical shape, the criteria of *ELG* < 2.54 and *CI* > 0.31 eliminated many elongated depression features that appeared to be stream channels, road ditches, and other man-made or natural features that were less likely to be sinkholes. Water-filled ponds usually have flat bottoms in the LiDAR DEM, resulting in depressions with low *STD*. Using the threshold of *STD* > 0.18 m, we removed water-filled ponds and other hydro features from consideration as potential sinkholes. Using these criterions, we were able to distinguish sinkholes from other non-sinkhole depressions. Some examples of non-sinkhole depressions are shown in Fig. 7.

After applying the sinkhole extraction criteria (area < 4000 m², depth > 0.5 m, *STD* > 0.18 m, *ELG* < 2.54, and *CI* > 0.31), the numbers of detected sinkholes of 1st, 2nd, 3rd, and 4th rank in the study area were 5299, 208, 37, and 10, respectively. Visual assessment of the results shows that the localized contour tree method is a very effective approach to identify sinkholes in the region. Fig. 8 shows some examples

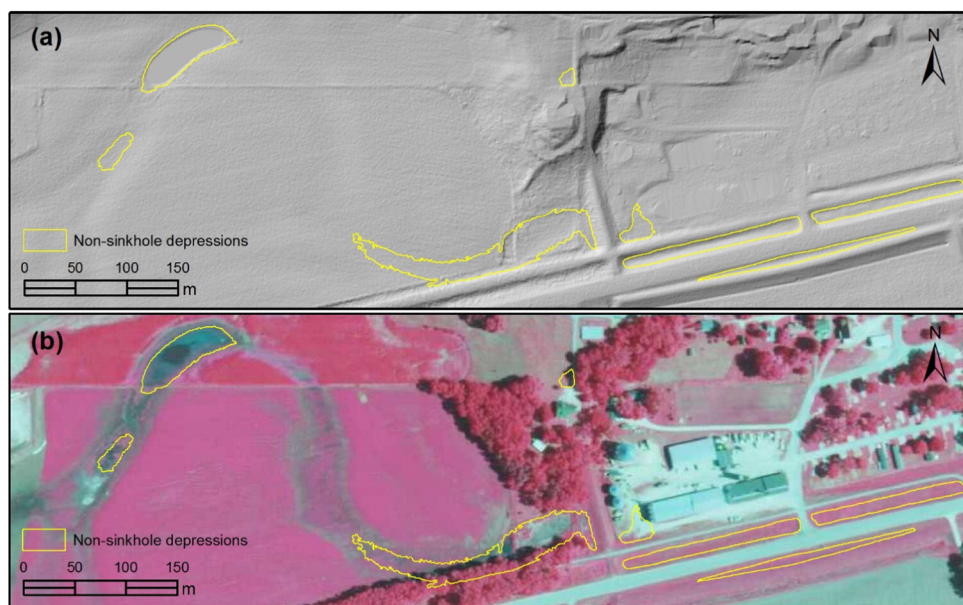


Fig. 7. Examples of non-sinkhole depression polygons overlain with LiDAR DEM shaded relief (a) or color infrared aerial imagery (b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

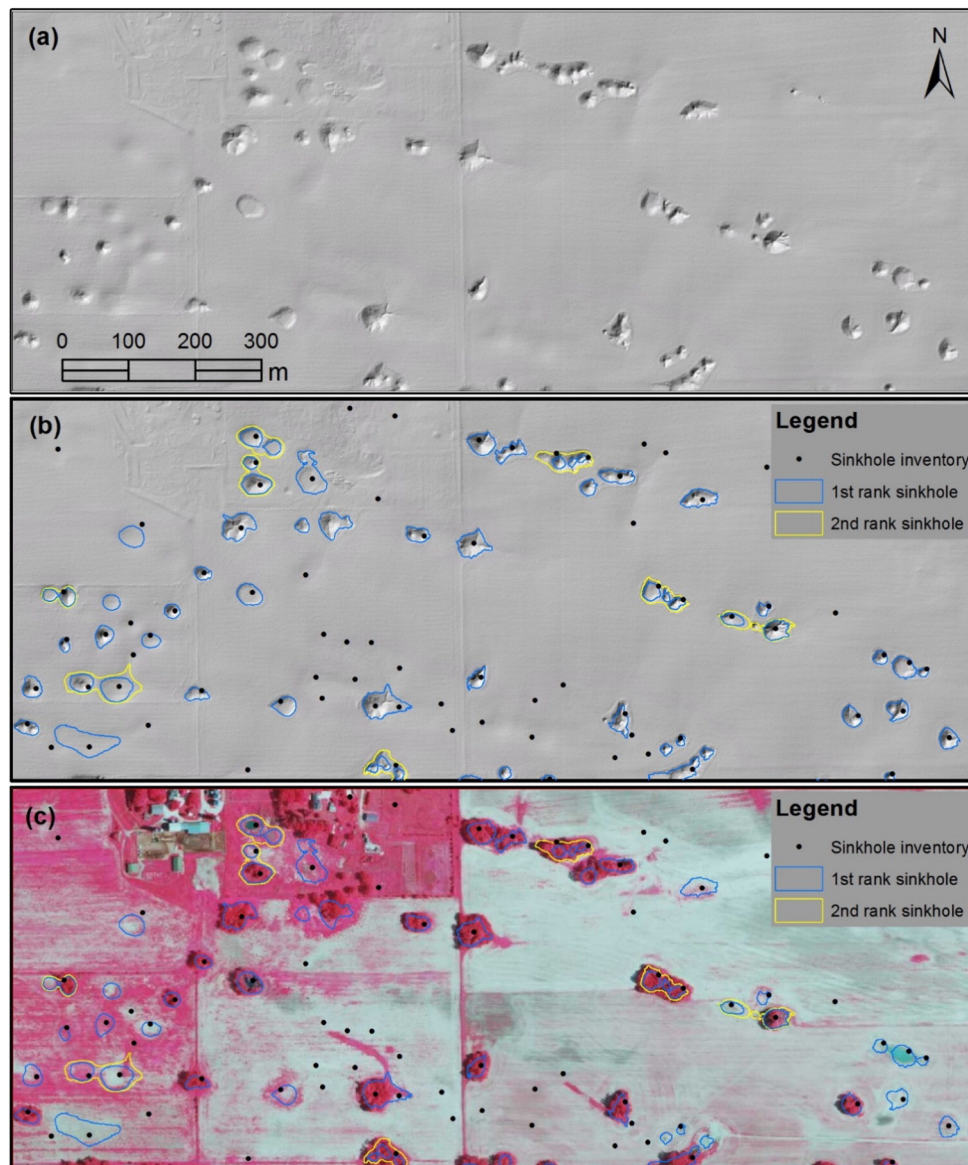


Fig. 8. LiDAR DEM shaded relief (a) and examples of extracted sinkhole boundaries overlain on LiDAR DEM shaded relief (b) and color infrared aerial imagery (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of extracted sinkhole boundaries overlain on the LiDAR DEM shaded relief and color infrared aerial imagery. Apparently, a large number of karst sinkholes were found under a dense forest or tree canopy. Those inventoried sinkholes not captured by the LiDAR data appeared to have been filled due to agricultural use or other reasons.

Table 2 provides a summary statistics of the extracted sinkholes at four different ranks. The identified 1st rank sinkholes have an aggregate area of 4.38 km², representing approximately 0.2% of the total land area of Fillmore County. The average area is 827 m² per sinkhole, which is the same as a circle with a diameter of 32.5 m. Median width and length of sinkholes increased with the sinkhole rank from 22.9 m at 1st rank to 125.4 m at 4th rank and from 35.5 m at 1st rank to 234.7 m at 4th rank, respectively. The median maximum depth of sinkhole ranged from 1.98 m at 1st rank to 5.69 m at 4th rank, while the deepest sinkhole at 4th rank reaching a depth of 8.18 m. The area of 1st rank sinkholes ranged from 100 to 3991 m², with a median size of 543.31 m². Following the classification of sinkholes according to Basso et al. (2013), 16.3% of 1st rank sinkholes were circular, 56.0% were elliptical, and 27.6% were elongated. Similarly, the lower rank sinkholes were much more

compact according to the compactness index with a gradual decrease from 1st rank to 4th rank.

Table 3 shows the comparison between inventoried sinkholes in the KFDB and sinkholes detected from the LiDAR DEM using the localized contour tree method. Out of the 6139 inventoried sinkholes in Fillmore County recorded in the KFDB, 1858 (30.2%) were successfully captured by the LiDAR data using the localized contour tree method for depression identification (see column “KFDB Detected” in Table 3). In order to evaluate the performance of LiDAR data for sinkhole mapping, we incorporated the National Land Cover Database (NLCD) 2011 into the analysis. The land cover types in Fillmore County are dominated by cultivated crops (41.5%), followed by pasture/hay (21.8%), deciduous forest (18.6%) and grassland/herbaceous (11.5%). Fig. 1 clearly shows that the majority of sinkholes were highly concentrated on flat hilltops between or adjacent to river valleys. The primary land cover type associated with these flat hilltop areas was cultivated crops. In other words, the majority of sinkholes were distributed in the agricultural land areas. A total of 2815 (45.9%) inventoried sinkholes in the KFDB were located within agricultural crops, followed by 1651 (26.9%) in grassland/herbaceous and

Table 2
Summary statistics of the extracted sinkholes of four different ranks.

Rank	Statistics	Width (m)	Length (m)	Area (m ²)	Volume (m ³)	Depth (m)	STD (m)	ELG	CI
1st rank (n = 5299)	Min	7.48	12.26	100.07	17.02	0.34	0.18	1.00	0.30
	5th percentile	10.86	16.27	126.64	54.91	0.75	0.20	1.10	0.37
	Median	22.88	35.54	543.31	419.77	1.98	0.53	1.51	0.67
	Mean	25.88	40.57	827.07	764.07	2.47	0.64	1.59	0.66
	95th percentile	51.53	85.29	2642.40	2563.55	5.66	1.44	2.32	0.91
	Max	79.79	140.15	3991.30	21,035.00	14.34	4.41	2.55	0.98
2nd rank (n = 208)	Min	15.08	33.83	395.50	341.47	1.36	0.29	1.10	0.02
	5th percentile	19.66	44.46	631.63	654.36	1.92	0.45	1.28	0.17
	Median	42.79	81.04	2023.11	2325.73	4.63	1.00	1.92	0.41
	Mean	50.07	100.94	3292.56	4524.89	4.61	1.01	2.07	0.41
	95th percentile	98.32	206.30	8610.03	8562.34	7.08	1.62	3.08	0.63
	Max	284.77	739.73	75,092.06	279,754.47	11.39	2.87	4.68	0.81
3rd rank (n = 37)	Min	36.94	71.90	1783.63	2410.65	2.10	0.40	1.20	0.02
	5th percentile	40.26	86.59	2168.55	2797.63	2.94	0.49	1.27	0.11
	Median	87.51	159.54	6827.78	7757.91	5.52	0.98	1.85	0.30
	Mean	106.07	195.36	10,371.91	11,410.61	5.34	1.01	1.95	0.29
	95th percentile	238.44	380.52	24,887.89	32,097.61	7.78	1.61	2.88	0.43
	Max	330.37	672.76	47,538.59	68,436.05	8.63	1.77	3.04	0.51
4th rank (n = 10)	Min	58.32	115.39	5191.55	5909.16	3.29	0.56	1.32	0.12
	5th percentile	67.31	134.64	5350.94	6428.07	3.38	0.64	1.36	0.12
	Median	125.38	234.68	12,933.67	12,973.62	5.69	1.04	1.49	0.27
	Mean	134.58	222.00	12,915.62	16,216.36	5.81	1.07	1.76	0.25
	95th percentile	199.79	279.65	22,518.36	33,052.69	7.78	1.60	2.68	0.33
	Max	210.59	280.28	25,995.62	39,412.51	8.18	1.75	2.71	0.34

849 (13.8%) in pasture/hay. Only 414 (6.7%) inventoried sinkholes were found to be located in deciduous forest.

On the contrary, 1007 (19.0%) out of 5299 1st rank sinkholes detected using the localized contour tree method were located in deciduous forest. Among these forested sinkholes, 891 were new sinkholes that had not been recorded in the KFDB. This dramatic increase can be attributed to the capability of LiDAR for penetrating through vegetation canopy, which enables us to map small karst features like sinkholes with much less interference from vegetation than aerial photography. However, only 1259 (23.8%) sinkholes were detected in agricultural areas, compared to 2815 in the KFDB. The possible explanation for this striking decline is that many sinkholes have been filled due to agricultural use or other man-made reasons. We found that 3645 (59.4%) out of 6139 sinkholes in the KFDB were not located in the depression areas of the LiDAR DEM. This indicates that more than half of the inventory sinkholes were no longer depression features, and could not be captured by LiDAR.

It should be noted that our localized contour tree method is fundamentally different from the previous raster-based methods for sinkhole mapping (Mukherjee, 2012; Zhu et al., 2014). Our method more explicitly derives geometric and topological properties of sinkholes. Fig. 9

illustrates the differences of sinkhole detection results between our localized contour tree method and the traditional sink-filling method. It is clear that many small sinkholes from the traditional method are nested within larger surface depressions. For example, the largest surface depression (area = 14,327 m²) in the center of Fig. 9a enclosed three smaller sinkholes inside. The sinkhole extraction criteria above (area < 4000 m²) would disqualify this largest depression from being considered as a potential sinkhole. Therefore, those three smaller sinkholes nested within the depression would also be eliminated accordingly using the sink-filling method. On the contrary, our contour tree method provides a much more accurate and reliable sinkhole mapping results (Fig. 9b).

5. Discussion and conclusions

In this study, we built on the work by Wu et al. (2015) who proposed the conceptual framework of the localized contour tree method for identifying surface depressions. We implemented the proposed conceptual framework and made the following improvements: (1) instead of using the whole DEM for the entire area to generate contours, we

Table 3
Comparison of the sinkhole inventory database and the LiDAR-derived sinkhole data (note: KFDB – sinkhole inventory points in Karst Feature Database; KFDB detected – sinkhole inventory points located within LiDAR-derived depressions using the contour tree method; KFDB not detected – sinkhole inventory points not located within LiDAR-derived depressions; and new sinkholes – sinkholes detected using the contour tree method but not recorded in KFDB).

Land cover types	Area (km ²)	Percentage	KFDB	KFDB detected	KFDB detected (%)	KFDB not detected	Contour tree	New sinkholes
Open water	4.6	0.2%	0	0	NA	0	0	0
Developed, open space	75.0	3.4%	273	95	34.8%	178	318	223
Developed, low intensity	39.9	1.8%	99	27	27.3%	72	187	160
Developed, medium intensity	4.3	0.2%	14	0	0.0%	14	16	16
Developed, high intensity	1.1	0.0%	3	0	0.0%	3	6	6
Barren land	0.7	0.0%	10	1	10.0%	9	5	4
Deciduous forest	416.4	18.6%	414	116	28.0%	298	1007	891
Evergreen forest	9.0	0.4%	3	1	33.3%	2	3	2
Mixed forest	0.5	0.0%	0	0	NA	0	1	1
Shrub/scrub	0.2	0.0%	0	0	NA	0	1	1
Grassland/herbaceous	256.2	11.5%	1651	602	36.5%	1049	1336	734
Pasture/hay	487.9	21.8%	849	250	29.4%	599	1125	875
Cultivated crops	926.2	41.5%	2815	764	27.1%	2051	1259	495
Woody wetlands	7.7	0.3%	6	2	33.3%	4	26	24
Emergent herbaceous wetlands	3.4	0.2%	2	0	0.0%	2	9	9
Total	2232.9	100.0%	6139	1858	30.3%	4281	5299	3441

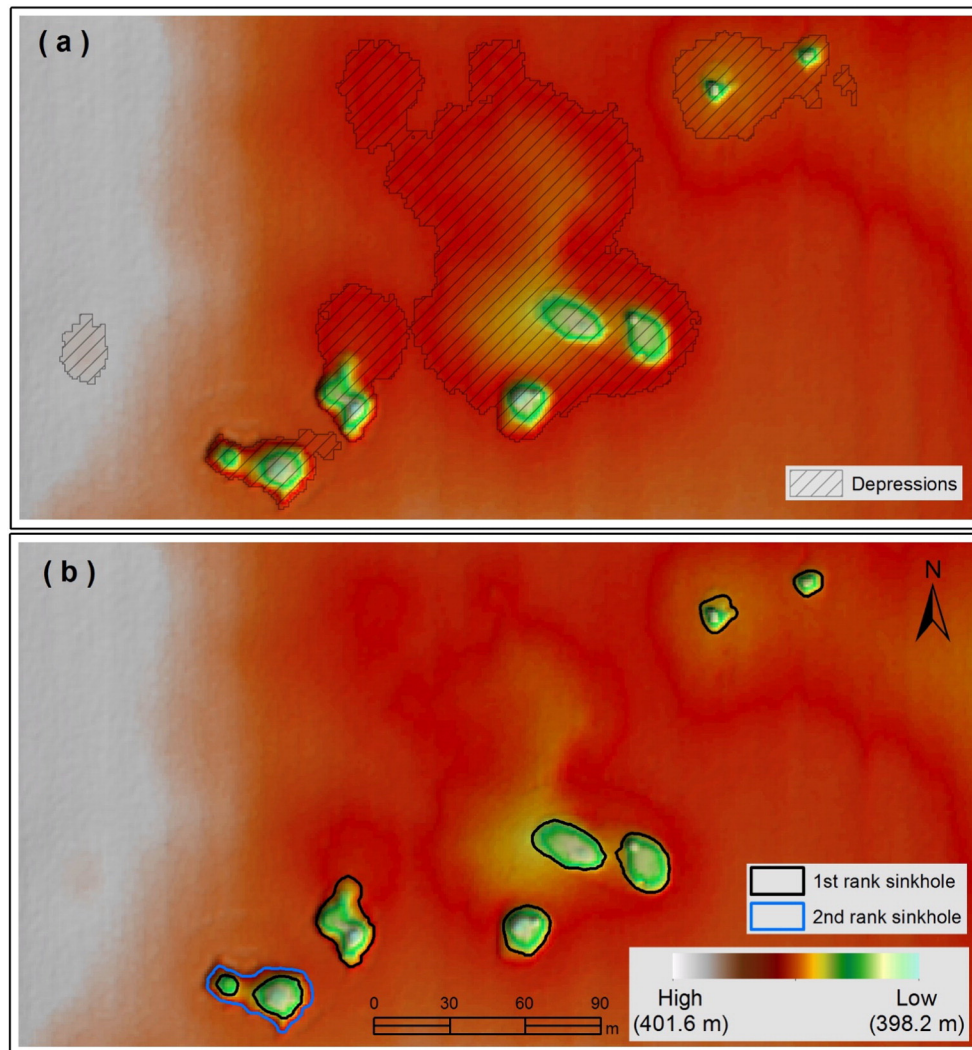


Fig 9. Sinkhole boundaries delineated using different methods. (a) The sink-filling method. (b) The localized contour tree method.

extracted a subset of the DEM representing depression areas, which can reduce the number of contours being generated and thus greatly reduce the computation time; (2) the computed morphometric properties of sinkholes provide critical information about sinkholes in addition to their locations; and (3) the improved algorithms have been implemented as an ArcGIS toolbox – Sinkhole Extraction Analyst. The automated procedure for detection of sinkholes described in our case study demonstrated its effectiveness and efficiency. In addition to detecting sinkhole locations, the localized contour tree method allowed for extracting sinkhole boundaries and quantifying sinkholes at different ranks across different spatial scales. Although we only used simple thresholding method to extract potential sinkholes from surface depressions, other machine learning-based methods such as decision tree and random forest could also be employed to facilitate sinkhole detection and susceptibility mapping, depending on the availability and quality of sinkhole training data.

It should be noted that the localized contour tree might not detect some shallow depressions whose depths are less than the contour interval, as these depressions might be absent from the contour maps even when these features actually exist in the landscape. This is the intrinsic limitation of DEM contouring. The number of artifact depressions resulted from LiDAR DEM error could be effectively reduced by setting appropriate thresholds of surface area and depth of depressions. However, there are no single values of area and depth thresholds that can achieve best results in all cases. As recommended by Li et al. (2011), soil and

climate conditions of a study site, the process of interest and the scope of the study all need to be taken into account when making the decision on selecting appropriate area and depth thresholds. Although this study only focused on potential sinkholes that were larger than 100 m², the method we proposed could be readily tailored to identify smaller sinkholes when finer-resolution DEMs become available.

It is also important to note that the LiDAR data flown during dry conditions on the ground is preferred for depression identification and sinkhole detection. LiDAR generally cannot penetrate water, meaning that topography of inundated depressions could not be captured by LiDAR data. This might result in true sinkholes covered with water not being detected if solely based on LiDAR data. Color infrared aerial photographs acquired in leaf-off conditions can facilitate the validation of sinkhole detection results when field verification is impractical. The acquisition date differences between LiDAR data and aerial photographs should be taken into account when sinkhole occurrences are not consistent between different data sets. Human development and land use practices, such as new residential development or agricultural expansion, could result in disappearance of sinkholes from the landscape. These new developments in the landscape might not be reflected in the LiDAR data or aerial photographs if they were acquired before these developments.

Mitigation of sinkhole related hazards in karst areas depends on locating and mapping these sinkholes accurately. High-resolution LiDAR data enables us to map small karst features like sinkholes with much less interference from vegetation than aerial photography. The

increasing availability of LiDAR data in the USA and some other countries would improve the precision of sinkhole mapping and result in a more accurate delineation of karst regions. The sinkhole mapping approach in this study may help to recognize previously unknown sinkholes and prevent human property losses. Although field verification is still an indispensable means for verifying sinkholes on the ground, many obvious sinkholes can be identified from LiDAR DEM. In addition, numerous non-sinkhole depressions can be identified from LiDAR data and eliminated from consideration, which could greatly reduce cost and field time.

Acknowledgements

We are grateful to the Minnesota Geospatial Commons, which provided data to support this research. Our appreciation also goes to two anonymous reviewers and Editor Dr. Takashi Oguchi for their helpful and useful comments and suggestions.

References

- Alexander, E.C., Maki, G.L., 1988. Sinkholes and Sinkhole Probability. *Geologic Atlas Olmsted County, Minnesota, County Atlas Series C-3, Plate 7(1)* p. 100,000.
- Al-Kouri, O., Al-Rawashdeh, S., Sadoun, B., Pradhan, B., 2013. Geospatial modeling for sinkholes hazard map based on GIS & RS data. *J. Geogr. Inf. Syst.* 5 (06), 584.
- Basso, A., Bruno, E., Parise, M., Pepe, M., 2013. Morphometric analysis of sinkholes in a karst coastal area of southern Apulia (Italy). *Environ. Earth Sci.* 70 (6), 2545–2559.
- Berardino, P., Fornaro, G., Lanari, R., Sansosti, E., 2002. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Trans. Geosci. Remote Sens.* 40 (11), 2375–2383.
- Chaudhuri, D., Samal, A., 2007. A simple method for fitting of bounding rectangle to closed regions. *Pattern Recogn.* 40 (7), 1981–1989.
- Davis, J.C., 2002. *Statistics and Data Analysis in Geology*. third ed. John Wiley & Sons, Inc., New York.
- Dinger, J., Zourarakis, D., Currens, J., 2007. Spectral enhancement and automated extraction of potential sinkhole features from NAIP imagery—initial investigations. *J. Environ. Inf.* 10 (1), 22–29.
- Doctor, D.H., Young, J.A., 2013. An Evaluation of Automated GIS Tools for Delineating Karst Sinkholes and Closed Depressions From 1-meter LiDAR-derived Digital Elevation Data.
- Galve, J.P., Castañeda, C., Gutiérrez, F., Herrera, G., 2015. Assessing sinkhole activity in the Ebro Valley mantled evaporite karst using advanced DInSAR. *Geomorphology* 229 (0), 30–44.
- Gao, Y., Alexander, E.C., 2003. A Mathematical Model for a Map of Relative Sinkhole Risk in Fillmore County, Minnesota, Sinkholes and the Engineering and Environmental Impacts of Karst (2003). ASCE, pp. 439–449.
- Gao, Y., Alexander, E.C., Tipping, R., 2002. The development of a karst feature database for southeastern Minnesota. *J. Cave Karst Stud.* 51–57.
- Gao, Y., Alexander, E.C., Barnes, R., 2005. Karst database implementation in Minnesota: analysis of sinkhole distribution. *Environ. Geol.* 47 (8), 1083–1098.
- Gutiérrez, F., Galve, J.P., Lucha, P., Castañeda, C., Bonachea, J., Guerrero, J., 2011. Integrating geomorphological mapping, trenching, InSAR and GPR for the identification and characterization of sinkholes: a review and application in the mantled evaporite karst of the Ebro Valley (NE Spain). *Geomorphology* 134 (1–2), 144–156.
- Gutiérrez, F., Parise, M., De Waele, J., Jourde, H., 2014. A review on natural and human-induced geohazards and impacts in karst. *Earth Sci. Rev.* 138 (0), 61–88.
- Heckmann, T., Schwanghart, W., Phillips, J.D., 2015. Graph theory—recent developments of its application in geomorphology. *Geomorphology* 243, 130–146.
- Hedges, J., Alexander, E., 1985. Karst-related features of the upper Mississippi Valley Region. *Stud. Speleol.* 6, 41–49.
- Huang, W., Deng, C.J., Day, M., 2014. Differentiating tower karst (fenglin) and cockpit karst (fengcong) using DEM contour, slope, and centroid. *Environ. Earth Sci.* 72 (2), 407–416.
- Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., Xian, G., 2013. A comprehensive change detection method for updating the national land cover database to circa 2011. *Remote Sens. Environ.* 132, 159–175.
- Kobal, M., Bertonecelj, I., Pirotti, F., Kutnar, L., 2014. Lidar processing for defining sinkhole characteristics under dense forest cover: a case study in the dinaric mountains, International Archives of the Photogrammetry. *Remote Sens. Spat. Inf. Sci. ISPRS Arch.* 113–118.
- Kweon, I.S., Kanade, T., 1994. Extracting topographic terrain features from elevation maps. *CVGIP: Image Underst.* 59 (2), 171–182.
- Larson, E.B., 2009. Advances in sinkhole mapping: a LiDAR survey of Houston County, Minnesota. 2009 Portland GSA Annual Meeting.
- Li, S., MacMillan, R.A., Lobb, D.A., McConkey, B.G., Moulin, A., Fraser, W.R., 2011. Lidar DEM error analyses and topographic depression identification in a hummocky landscape in the prairie region of Canada. *Geomorphology* 129 (3–4), 263–275.
- Lindsay, J.B., 2014. The Whitebox Geospatial Analysis Tools Project and Open-access GIS. *Proceedings of the GIS Research UK 22nd Annual Conference*. The University of Glasgow.
- Lindsay, J.B., Creed, I.F., 2006. Distinguishing actual and artefact depressions in digital elevation data. *Comput. Geosci.* 32 (8), 1192–1204.
- Liu, H., Wang, L., Sherman, D., Gao, Y., Wu, Q., 2010. An object-based conceptual framework and computational method for representing and analyzing coastal morphological changes. *Int. J. Geogr. Inf. Sci.* 24 (7), 1015–1041.
- Magdalene, S.C.C., Alexander, E., 1995. Sinkhole Distribution in Winona County, Minnesota, Revisited. University of Minnesota.
- Miao, X., Qiu, X., Wu, S.-S., Luo, J., Gouzie, D.R., Xie, H., 2013. Developing efficient procedures for automated sinkhole extraction from Lidar DEMs. *Photogramm. Eng. Remote Sens.* 79 (6), 545–554.
- Minnesota Geospatial Commons, 2005. Karst feature inventory points. (Available online) <https://gisdata.mn.gov/dataset/geos-karst-feature-inventory-pts> (accessed May 3, 2016).
- Minnesota Geospatial Commons, 2008. LiDAR elevation, Southeast Minnesota, 2008. (Available online) <https://gisdata.mn.gov/dataset/elev-lidar-semn2008> (accessed May 3, 2016).
- Minnesota Geospatial Commons, 2015. Digital orthoimagery, southern Minnesota, spring 2011; 0.5-m resolution. Available online: <https://gisdata.mn.gov/dataset/base-saip-so2011-50cm-airphotos> (accessed May 3, 2016).
- Montane, J.M., 2001. Geophysical Analysis of a Central Florida Karst Terrain Using Light Detection and Ranging (LiDAR) and Ground Penetrating Radar (GPR) Derived Surfaces.
- Mukherjee, A., 2012. GIS analysis of sinkhole distribution in Nixa, Missouri. 2012 GSA Annual Meeting in Charlotte.
- Nwokebuihe, S.C., Torgashov, E.V., Anderson, N., 2014. A description of an effective sinkhole investigation approach: a case study of a site in Greene County, Missouri. 27th Annual Symposium on the Application of Geophysics to Engineering and Environmental Problems (SAGEEP).
- Obu, J., Podobnikar, T., 2013. Algorithm for karst depression recognition using digital terrain models. *Geodetski Vestnik* 57 (2), 260–270.
- Phillips, J.D., Schwanghart, W., Heckmann, T., 2015. Graph theory in the geosciences. *Earth Sci. Rev.* 143, 147–160.
- Pratt, W.K., 1991. *Digital Image Processing*. second ed. John Wiley & Sons, Inc., New York.
- Rahimi, M., Alexander, E.C., 2013. Locating Sinkholes in LiDAR Coverage of a Glacio-fluvial Karst, Winona County, MN.
- Seale, L.D., Florea, L.J., Vacher, H.L., Brinkmann, R., 2008. Using ALSM to map sinkholes in the urbanized covered karst of Pinellas County, Florida—1, methodological considerations. *Environ. Geol.* 54 (5), 995–1005.
- Shaban, A., Darwich, T., 2011. The role of sinkholes in groundwater recharge in the high mountains of Lebanon. *J. Environ. Hydrol.* 19.
- Shade, B., Alexander, S., Alexander, E., Martin, S., 2001. Sinkhole Distribution, Depth to Bedrock, and Bedrock Topography. *Geologic Atlas of Pine County, Minnesota, County Atlas Series C-13, Part a, Plate 6(1)* p. 100,000.
- Shukunobe, M., 2012. Using LiDAR Data to Map Depressions and Sinkholes in Mammoth Cave National Park.
- Taheri, K., Gutiérrez, F., Mohseni, H., Raeisi, E., Taheri, M., 2015. Sinkhole susceptibility mapping using the analytical hierarchy process (AHP) and magnitude–frequency relationships: a case study in Hamadan province, Iran. *Geomorphology* 234 (0), 64–79.
- Vacher, H.L., Seale, L.D., Florea, L.J., Brinkmann, R., 2008. Using ALSM to map sinkholes in the urbanized covered karst of Pinellas County, Florida—2. Accuracy statistics. *Environ. Geol.* 54 (5), 1007–1015.
- Wang, L., Liu, H., 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *Int. J. Geogr. Inf. Sci.* 20 (2), 193–213.
- Witthuhn, M., Alexander, E., 1995. Sinkholes and Sinkhole Probability. *Geologic Atlas Fillmore County, Minnesota, County Atlas Series C-8, Part B, Plate 8(1)* p. 100,000.
- Wu, Q., 2011. Object-oriented Representation and Analysis of Coastal Changes for Hurricane-induced Damage Assessment (MA Electronic Thesis or Dissertation) University of Cincinnati (30 pp.).
- Wu, Q., Lane, C., Liu, H., 2014. An effective method for detecting potential woodland vernal pools using high-resolution LiDAR data and aerial imagery. *Remote Sens.* 6 (11), 11444–11467.
- Wu, Q., Liu, H., Wang, S., Yu, B., Beck, R., Hinkel, K., 2015. A localized contour tree method for deriving geometric and topological properties of complex surface depressions based on high-resolution topographical data. *Int. J. Geogr. Inf. Sci.* 29 (12), 2041–2060.
- Zhu, J., Taylor, T.P., Currens, J.C., Crawford, M.M., 2014. Improved karst sinkhole mapping in Kentucky using LiDAR techniques: a pilot study in Floyds Fork watershed. *J. Cave Karst Stud.* 76 (3), 207–216.