



Transportation Consortium of South-Central States

Solving Emerging Transportation Resiliency, Sustainability, and Economic Challenges through the Use of Innovative Materials and Construction Methods: From Research to Implementation

Karst Sinkhole Detecting and Mapping Using Airborne LiDAR

Project No. 18GTUNM01

Lead University: University of New Mexico

Final Report
August 2019

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgements

The authors would like express our greatest gratitude to the students who participated in this research projects. The authors would also like to thank Tim Parker from the New Mexico Department of Transportation District 2 and George Veni from the National Cave and Karst Research Institute. The authors would also like to thank Dr. Shawn Penman, Mr. Claude Morelli, and Dr. David Hadwiger for serving as the Project Review Committee members.

TECHNICAL DOCUMENTATION PAGE

1. Project No. 18GTUNM01	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Karst Sinkhole Detecting and Mapping Using Airborne LiDAR		5. Report Date Aug. 2019	
7. Author(s) PI: Su Zhang https://orcid.org/0000-0002-0396-2518 Co-PI: Shirley V. Baros https://orcid.org/0000-0002-0842-1219 Co-PI: Susan Bogus Halter https://orcid.org/0000-0003-1685-6367		6. Performing Organization Code	
9. Performing Organization Name and Address Transportation Consortium of South-Central States (Tran-SET) University Transportation Center for Region 6 3319 Patrick F. Taylor Hall, Louisiana State University, Baton Rouge, LA 70803		8. Performing Organization Report No.	
12. Sponsoring Agency Name and Address New Mexico Department of Transportation (NMDOT) District 2 1120 Cerrillos Road, Santa Fe, NM 87504		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. 69A3551747106	
		13. Type of Report and Period Covered Final Research Report Mar. 2018 – Mar. 2019	
		14. Sponsoring Agency Code	
15. Supplementary Notes Report uploaded and accessible at Tran-SET's website (http://transet.lsu.edu/) .			
16. Abstract <p>The focus of this study is to detect sinkhole hazards using airborne light detection and ranging (LiDAR) data. The premise is sinkholes, particularly those close to transportation infrastructure assets, could cause substantial damages to infrastructure assets, and therefore, being able to accurately and rapidly detect them is essential. However, it is expensive, time-consuming, labor-intensive, and unsafe to survey sinkholes using conventional ground observation methods. This research project was focused on developing accurate and rapid airborne LiDAR-based sinkhole detection and mapping methods, and transfer the technologies to transportation engineers for implementation and workforce development. The project team also identified best practices for implementation of a state-level sinkhole hazard management system (SHMS). In addition, a guidebook was developed for airborne LiDAR-based sinkhole detection and mapping for professional education and training.</p> <p>The effectiveness of LiDAR to detect existing sinkholes has received very limited attention. Most of the research on LiDAR-based sinkhole detection postulates that morphological-based surface feature extraction methods can effectively detect sinkholes because of their geometric properties – sinkholes are oval-shaped concave depressions in the Earth's surface. However, sinkholes have varying sizes, shapes, and appearance given various landforms, which adds even greater challenges to further improving the detection accuracy of methods that are based solely on morphology; for example, a dry stock pond may be incorrectly detected as a sinkhole. The proposed research used airborne LiDAR data in combination with auxiliary context such as site and association to improve the accuracy of the morphological-based sinkhole detection methods, and implement these by developing tools that can be used in standard geographic information systems (GIS). This methodology allows for the development of a robust LiDAR-based sinkhole detection toolset that provides an adequate degree of accuracy while maximizing the ability to assist inspectors with varying expertise.</p>			
17. Key Words Airborne LiDAR, GIS, Sinkhole Detection, Sinkhole Mapping		18. Distribution Statement No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 39	22. Price

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
AREA				
in ²	square inches	645.2	square millimeters	mm ²
ft ²	square feet	0.093	square meters	m ²
yd ²	square yard	0.836	square meters	m ²
ac	acres	0.405	hectares	ha
mi ²	square miles	2.59	square kilometers	km ²
VOLUME				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft ³	cubic feet	0.028	cubic meters	m ³
yd ³	cubic yards	0.765	cubic meters	m ³
NOTE: volumes greater than 1000 L shall be shown in m ³				
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
TEMPERATURE (exact degrees)				
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m ²	cd/m ²
FORCE and PRESSURE or STRESS				
lbf	poundforce	4.45	newtons	N
lbf/in ²	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
AREA				
mm ²	square millimeters	0.0016	square inches	in ²
m ²	square meters	10.764	square feet	ft ²
m ²	square meters	1.195	square yards	yd ²
ha	hectares	2.47	acres	ac
km ²	square kilometers	0.386	square miles	mi ²
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m ³	cubic meters	35.314	cubic feet	ft ³
m ³	cubic meters	1.307	cubic yards	yd ³
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
TEMPERATURE (exact degrees)				
°C	Celsius	1.8C+32	Fahrenheit	°F
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m ²	candela/m ²	0.2919	foot-Lamberts	fl
FORCE and PRESSURE or STRESS				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in ²

TABLE OF CONTENTS

TECHNICAL DOCUMENTATION PAGE	ii
TABLE OF CONTENTS.....	iv
LIST OF FIGURES	vi
LIST OF TABLES	viii
ACRONYMS, ABBREVIATIONS, AND SYMBOLS	ix
EXECUTIVE SUMMARY	x
1. INTRODUCTION	1
2. OBJECTIVES.....	3
3. LITERATURE REVIEW	4
4. METHODOLOGY	7
4.1. Data Acquisition and Preparation	7
4.2. Preliminary Sinkhole Detection and Map Development	12
4.3. Final Sinkhole Detection and Map Development.....	14
4.3.1. Context-based Feature Extraction.....	14
4.3.2. Morphology-based Feature Extraction.....	17
4.3.3. Methodology Workflow Summary	18
4.4. Sinkhole Map Validation	19
5. ANALYSIS AND FINDINGS	20
5.1. Data Preparation Results.....	20
5.2. Preliminary Sinkhole Detection and Mapping Results.....	22
5.3. Final Sinkhole Detection and Mapping Results	24
5.4. Sinkhole Map Validation	32
5.4.1. Overall Level Validation.....	32
5.4.2. Individual Level Validation	34
6. CONCLUSIONS.....	38
REFERENCES	40
APPENDIX A: BEST PRACTICES FOR IMPLEMENTING A STATE-LEVEL SINKHOLE HAZARD MANAGEMENT SYSTEM (SHMS).....	42

APPENDIX B: LIDAR BASED SINKHOLE DETECTION AND MAPPING TOOLSET
GUIDEBOOK..... 43

APPENDIX C. LIDAR BASED SINKHOLE DETECTION AND MAPPING TOOLSET CODES
..... 56

LIST OF FIGURES

Figure 1. Roswell-Carlsbad District and Rio Hondo and Upper Pecos-Black Watershed, New Mexico.	7
Figure 2. Signs warn drivers to be wary of potential sinkholes near Carlsbad (adopted from Carlsbad Current Argus).	8
Figure 3. Aerial photo coverage for Carlsbad and Loving Village (adopted from NMDOT).	8
Figure 4. The display of a point cloud that covers a small portion of the City of Albuquerque. ...	9
Figure 5. The Rio Hondo Watershed boundary (dark red) and individual LAS tiles (light red). ...	9
Figure 6. An illustration of the difference between DTM and DSM (adopted from 3DMetrica). ...	10
Figure 7. The processed DEMs tile for the study area; 40 tiles with a total area of 90 km ²	11
Figure 8. Sinkhole detection principle of identifying the distinct shape-transition boundary in reference to the surrounding terrain (adopted from NCKRI).	12
Figure 9. Eight flow directions for all neighboring cells of a focal cell.	12
Figure 10. An illustration of the process of filling sinks.	13
Figure 11. The Interface of the geoprocessing tool for preliminary sinkhole detection.	13
Figure 12. The toolbox that can be added to the ArcGIS software interface.	13
Figure 13. An illustration of remote sensing imagery classification (adopted from Land Info). .	14
Figure 14. Pixel-based image analysis classification vs. OBIA classification.	15
Figure 15. An illustration of the Mean Shift segmentation (adopted from chioka.in).	16
Figure 16. Flowchart of the methodology.	18
Figure 17. The Sinkhole Toolbox that is compatible with ArcGIS.	19
Figure 18. The user interface of the DEM Creator tool that is compatible with ArcGIS.	20
Figure 19. The mosaicked DEM layer when using multiple LAS tiles as input.	21
Figure 20. The individual DEM layers when using a single LAS tile as input.	22
Figure 21. The output of the preliminary sinkhole detection tool.	23
Figure 22. The user interface of the Sinkhole Extractor tool that is compatible with ArcGIS.	24
Figure 23. Sinkhole detection result when using 32-Bit Float with minimum 100 m ² and maximum 5,000,000 m ² area limit.	26
Figure 24. Sinkhole detection result when using 32-Bit Signed with minimum 100 m ² and maximum 5,000,000 m ² area limit.	27
Figure 25. A false positively detected sinkhole which is actually a reservoir.	28

Figure 26. Sinkhole detection result when using 32-Bit Signed with minimum 100 m ² and maximum 30,000 m ² area limit.	29
Figure 27. The location of the extract mask and the study area.	30
Figure 28. Sinkhole detection result when using 32-Bit Signed with minimum 100 m ² and maximum 30,000 m ² area limit and with the extract mask.	31
Figure 29. The zoomed in look of the area that is highlighted in Figure 28.	32
Figure 30. The distribution of the airborne LiDAR detected sinkhole surface area values.	34
Figure 31. The distribution of the ground-truth sinkhole surface area values.	34
Figure 32. Box plot of the airborne LiDAR detected sinkhole and ground-truth sinkhole morphometric measurements.	35
Figure 33. The boundary of LiDAR detected sinkholes (in beige color) and the ground-truth sinkholes (in jade color).	36

LIST OF TABLES

Table 1. Context information elements and matching application.	17
Table 2. Bit depth options and their applications.	25
Table 3. Cohen’s Kappa matrix.	32
Table 4. Cohen’s Kappa Matrix for LiDAR detected sinkholes.....	33
Table 5. Cohen’s Kappa coefficient interpretation.	33
Table 6. Normality test results for LiDAR detected and ground-truth sinkhole layers.....	35

ACRONYMS, ABBREVIATIONS, AND SYMBOLS

DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model.
EDAC	Earth Data Analysis Center
FEMA	Federal Emergency Management Agency
GIS	Geographic Information Systems
LiDAR	Light Detection and Ranging
NCKRI	National Cave and Karst Research Institute
NMDOT	New Mexico Department of Transportation
OBIA	Object Based Image Analysis
UNM	University of New Mexico
USGS	United States Geological Survey

EXECUTIVE SUMMARY

Sinkhole subsidence and collapse may cause significant structural damage to transportation infrastructure assets and are a major problem in karst areas where soluble bedrock is predominant. Subsequently, transportation infrastructure management agencies at all levels (e.g., federal, state, tribal, and local) dedicate a large amount of time and money to routinely detect and map sinkholes as part of their infrastructure asset management programs. These collected sinkhole data are used by the aforementioned agencies to determine the extent and severity of sinkholes, and then to make decisions on the distribution of limited resources for sinkhole hazard mitigation to avoid potential public safety threats and problems.

Sinkhole subsidence or collapse is a natural hazard of national scope – sinkholes are found in all 50 states, but predominantly in areas with karst terrain. Along with recent years' economic growth, man-made transportation infrastructure assets have expanded onto karst terrain that was previously rural and sparsely developed. Sinkholes pose a severe threat to these newly built transportation infrastructure assets. Therefore, being able to accurately detect and map existing sinkholes is very important for transportation infrastructure sinkhole risk assessment and hazard mitigation, and ultimately, leading to more informed planning of transportation infrastructure and public safety.

Traditionally, sinkholes are primarily detected through area reconnaissance, which includes visual inspection of a site to locate existing sinkholes or device inspection of a site to locate potential sinkholes or previously filled sinkholes. This method is expensive, time-consuming, labor-intensive, and potentially dangerous to the inspectors. More importantly, because of its ground-based nature, this method is limited not only by the accessibility of the site but also the ability of the inspectors or devices to observe the entire site. For example, heavy-vegetation in a certain site may make it extremely difficult or even impossible to conduct any reconnaissance at all.

To overcome these challenges, researchers from the Earth Data Analysis Center (EDAC) and the Department of Civil, Construction, and Environmental Engineering at the University of New Mexico (UNM) collaborated on this research to explore the utility of airborne Light Detection and Ranging (LiDAR) in detecting and mapping sinkholes. The research team used airborne LiDAR data in combination with not only object-based image analysis (OBIA) techniques but also auxiliary context information such as site and association to improve the accuracy of the current morphology-based sinkhole detection methods. Specifically, the primary input data for this research are airborne LiDAR, which uses laser light to densely sample the Earth's surface to produce highly accurately measurements in x, y, and z dimensions. In addition, this research used OBIA techniques to delineate the sinkhole boundaries. Additionally, auxiliary context information such as site (e.g., specific terrain does not have sinkholes) and associate (e.g., urban vs. rural) were used to further improve the accuracy of sinkhole detection and mapping.

This research also developed a robust toolset that be used in standard geographic information systems (GIS) for operational implementation. This toolset can be used to detect and map sinkholes with an adequate degree of accuracy while maximizing the ability to assist inspectors with varying expertise. Results revealed that airborne LiDAR detected sinkholes and ground surveyed sinkholes have statistically similar morphometric measurements. This research also developed a guidebook on using the developed tools for professional education and training. The ultimate goal of this research is to train a new generation of transportation engineers that can effectively use the developed techniques and tools to accurately and rapidly detect and map sinkholes.

1. INTRODUCTION

Sinkholes are natural depressions in the Earth's surface that form in places where carbonate rocks are dissolved from water and overlying soil particles are carried away underground, causing the surface to subside gradually or collapse suddenly (1,2,3). Sinkholes are a major geohazard in many karst areas because they can destroy infrastructure assets such as roadways and buildings. The Cambridge English Dictionary defines karst as an area of land formed of rock such as limestone that is worn away by water to make caves and other formations (4). Sinkholes have also played an important role in groundwater systems by providing connections between surface water and groundwater – collecting surface precipitation and draining it internally into the subsurface (3).

In general, sinkholes can be classified into three types, including dissolution sinkholes, cover-subsidence sinkholes, and cover-collapse sinkholes (5). For dissolution sinkholes, small amounts of soil or vegetation are found over the limestone or other bedrock, and water from precipitation and runoff slowly trickles through cracks and crevices in the bedrock to dissolve it (6). Therefore, dissolution sinkholes gradually form in the Earth's surface. Cover-subsidence sinkholes occur in areas sand covers the bedrock. The sand filters down into openings in the rock, causing the land surface to sink (6). Similar to dissolution sinkholes, cover-subsidence sinkholes occur slowly over time. For cover-collapse sinkholes, ground sediments gradually erode or spall into the cavern from the bottom, which will cause the ground suddenly crumble. When the thin layer between the surface and the underground openings collapse, a sinkhole will occur and swallow any objects (e.g., trees and cars) or structures (e.g., buildings and roads) above (6).

Both sinkhole sudden collapse and gradual subsidence may cause significant damage to infrastructure assets such as roadways, bridges, and buildings (2). In rural areas, sinkholes could cause the loss of arable land as well as infrastructure assets such as roadways and bridges. In an urban setting sinkholes also pose a threat to humans and area of particular concern to urban planners, homeowners, and insurance companies (2,7,8). In recent years, sinkhole risks are becoming particularly severe in urban areas that lack careful planning and where land depressions are frequently reported and developed (9).

Sinkhole subsidence and collapse is a natural hazard of national scope – sinkholes are found in all 50 states, but predominantly in areas with karst terrain. A substantial amount of karst terrain is located in Transportation Consortium of South-Central States (Tran-SET) states, including Louisiana, Oklahoma, Texas, Arkansas, and New Mexico (10). Along with recent years' economic growth, man-made transportation infrastructure and buildings have expanded onto karst terrain that was previously rural and sparsely developed (11). Sinkholes pose a severe threat to these newly built infrastructure assets – sinkhole hazards could cause substantial structural damage. Repair of sinkhole damages to buildings, highways, and other infrastructure systems represents a significant national cost (11). For example, sparse and incomplete data show that the average cost of karst-related damages in the United States over the last 15 years is estimated to be at least \$300 million per year and the actual cost is probably higher (11). Because of their sudden appearance and hazardous nature, it is important to characterize sinkholes' embryonic structure and identify their location at an early stage (12).

Therefore, being able to accurately detect and map existing sinkholes (sinkholes that have occurred and appeared on the ground) is critical for transportation infrastructure sinkhole risk assessment and hazard mitigation, and ultimately, leading to more informed planning of transportation infrastructure and public safety. However, the current methods for detecting and mapping existing

sinkholes are expensive, time-consuming, labor-intensive, tedious, subjective, potentially dangerous to inspectors, requiring specialized staff on a regular basis, and can exhibit a high degree of variability, thereby causing inconsistencies in surveyed data over space and across evaluation. In addition, sinkhole data collected by current methods can only be used for a single purpose and cannot be shared with other government agencies (e.g., the U.S. Geological Survey) to reduce the cost. To solve these problems, this research is focused on developing an accurate and rapid airborne LiDAR-based existing sinkhole detection and mapping method and transferring the technologies to transportation management agencies for implementation and workforce development via a developed toolset. More specifically, this research is focused on developing an airborne LiDAR-based sinkhole detection and mapping toolset which will enable transportation engineers to rapidly detect and map sinkholes with an adequate degree of accuracy while maximizing the ability to assist transportation engineers with varying expertise.

2. OBJECTIVES

The main objective of this research is to develop a new approach to accurately and rapidly detect and map existing sinkholes through the use of airborne LiDAR data and transfer the technologies to transportation infrastructure management agencies at all levels (e.g., federal, state, tribal, and local) for implementation and workforce development. Specifically, this research project will: (1) develop a complete process and toolset for detecting and mapping collapse sinkholes through the use of airborne LiDAR data; (2) identify best practices for the effective implementation of a statewide sinkhole hazard management system (SHMS); and (3) develop a guidebook for airborne LiDAR-based existing sinkhole detection and mapping which can be used for professional education and training.

3. LITERATURE REVIEW

As mentioned in the introduction section, sinkholes, particularly those close to infrastructure assets, could cause substantial structural damage to infrastructure assets, and therefore, being able to accurately and rapidly detect and map them is vital. Subsequently, transportation infrastructure management agencies at all levels (e.g., federal, state, tribal, and local) dedicate a large amount of time and money to detect and map existing sinkholes as part of their infrastructure asset management programs. These collected sinkhole data are used by the aforementioned agencies to determine the extent and severity of sinkholes, and then to make decisions on the distribution of limited resources for sinkhole hazard mitigation to avoid potential public safety threats and problems.

Traditionally, sinkholes are detected through area reconnaissance, which includes visual inspection of a site to locate existing sinkholes or instrumental inspection of site to locate potential sinkholes or previously filled sinkholes. For visual inspection, collapse sinkhole data are collected by inspector walking along a sinkhole and rating the risk factors. These data are primarily handwritten data and attached to archived images acquired by inspectors on the ground. Instrumental inspection involves using instrument for locating potential sinkholes, and the most popular instruments include resistivity and electromagnetic survey equipment, seismic survey equipment, ground penetration radar (GPR), microgravity survey equipment, probing and boring equipment, and video televiewer (13). Both visual inspection and instrumental inspection are expensive, time-consuming, labor-intensive, and potentially dangerous to inspectors. Visual inspection is also subjective and data collected by different inspectors can exhibit a high degree of variability which causes inconsistencies in surveyed data over space and across evaluation. Instrumental inspection is objective, but it requires specialized staff on a regular basis. Because of its high price and limited suitability (potential sinkholes or previously filled sinkholes), it has not been used for routine sinkhole inspection, but instead, it has been used for on-call or emergency inspection. More importantly, because of its ground-based nature, area reconnaissance method is limited not only by the accessibility of the site but also the ability of the inspectors or instruments to observe the entire site. For example, heavy vegetation may make it difficult or even impossible to inspect certain sites (13).

Another method for sinkhole detection is through a review of topographic maps, contour maps, geologic maps, and sinkhole inventory maps. These maps provide fair to good reliability for locating existing sinkholes and for determining the susceptibility of an area to future sinkhole development due to geologic and hydrologic factors (13). However, the accuracy of detecting sinkholes from these maps depends on sinkhole size, map scale, contour interval, and slope of the ground surface (14). For example, on a U.S. Geological Survey (USGS) 1:24,000 scale 7.5 minute quadrangle map, the minimum size of detectable sinkholes is 48 m in diameter, which fails to detect any sinkholes that are smaller than that dimension, although small sinkholes could be far more dangerous than large sinkholes (2). Therefore, this method has not been operationally used as a primary method for detecting and mapping existing sinkholes, but instead, as a complementary method.

Other researchers have explored the utility of imagery collected by optical sensors in detecting and mapping existing sinkholes. These studies relied on large-scale and high-spatial resolution aerial photos and satellite images – locating sinkholes in the images and then digitizing them in. For example, Wilson (15) identified 900 existing sinkholes from the 1926 and 1995 aerial photos

collected for the Pinellas County in the State of Florida. Ernst et al. (16) used QuickBird and WorldView satellite images to study the regional variation of sinkhole morphology. Examination of stereoscopic aerial photos may also be used to detect and locate existing sinkholes. However, the reliability of the use of aerial photos or satellite images for detecting and mapping sinkholes depends on the scale of the images and the experience of the observers (13). This method is prone to observers' subjective bias, and it is also expensive to deploy, requiring specially trained observers and substantial technical expertise, which ultimately prevents the further exploration of their applications in detecting and mapping existing sinkholes.

In recent years, LiDAR technology has been developed for detailed Earth's surface elevation data collection. LiDAR uses laser light to densely sample the Earth's surface to produce highly accurate measurements in x, y, and z dimensions. There are two types of LiDAR, including airborne LiDAR and terrestrial LiDAR. For airborne LiDAR, the data collection system is installed on either a helicopter or fixed-wing airplane. For terrestrial LiDAR, the data collection system is installed on a moving platform (e.g., a vehicle) or static platform (e.g., a tripod). Currently, airborne LiDAR is primarily used for aerial mapping applications. Compared with traditional optical sensor based remote sensing (e.g., stereoscopic aerial photography and satellite imagery), airborne LiDAR can collect elevation data with much higher accuracy (17), allowing the examination of the Earth's surface evaluation change accurately and rapidly. In addition, LiDAR has a unique capability – it can filter through vegetation to the ground as long as light can be seen under vegetation canopy. These features make airborne LiDAR hold the potential to be used as an effective means to detect and analyze any natural hazards characterized with surface depression such as sinkholes. In addition, many states in the United States is in the process of acquiring nationwide airborne LiDAR data (USGS Quality Level 2) as part the national 3D Elevation Program (3DEP). This program is a collaborative effort among government entities, academia, and the private sectors to collect high-spatial resolution 3D data across the United States (18). This project is expected to deliver 1 m (5 m for Alaska) digital elevation models (DEMs) with metadata and accuracy reports. High-quality LiDAR collected by the 3DEP will be provided to the public for free access and use. Coupled with advanced feature extraction techniques in remote sensing, these high-quality data offer a great opportunity to detect and map existing sinkholes in a rapid, accurate, and cost-effective manner on an unprecedented national scale.

Airborne LiDAR-based sinkhole detection techniques have been the subject of previous research and several methods have been developed and investigated (2,8,12,19,20,21,22). However, the effectiveness of using airborne LiDAR to detect and map existing sinkholes has received very limited attention. Most of the aforementioned research on airborne LiDAR-based sinkhole detection and mapping postulated that morphology-based surface feature extraction can effectively detect existing sinkholes because geometrically sinkholes are oval-shaped concave depression in the Earth's surface.

Therefore, the majority of previous studies focused on detecting objects that form a distinct shape-transition in reference to their surroundings and area embedded within the terrain. However, these methods have limited accuracy in detecting collapse sinkholes because they only consider sinkholes' morphology characteristics. The reality is that sinkholes have varying sizes, shapes, and appearance under various landforms (12), and they may not even exist in certain contexts. For example, a dry stock pond may be false positively detected as a sinkhole. In addition, these aforementioned research did not develop any functional tools to assist with their operational deployment, and in order to realize their intended function, specialized staff members are required

on a regular basis to carry out the complicated algorithms. That being said, only the developers of these algorithms know how to use them to detect and map sinkholes from airborne LiDAR.

When trying to identify sinkholes from airborne LiDAR data, several issues ought to be considered: (1) the quality of the LiDAR data, including spatial resolution and vertical accuracy – government entities collect LiDAR data with different quality levels; (2) the skills of the sinkhole inspectors – this reflects if the inspectors have specialized skills to use LiDAR data to conduct sinkhole detection and mapping; and (3) the usefulness of sharing tools – for the benefit of researchers or engineers or the benefit of the public at large. The objective of sinkhole detection in general is to find a balance between the quality of LiDAR data and the skills of sinkhole inspectors and the usefulness of tool sharing.

This research used airborne LiDAR data in combination with not only object-based image analysis (OBIA) techniques but also auxiliary context information such as site (e.g., specific terrain does not have sinkholes) and associate (e.g., urban vs. rural) to improve the accuracy of the current morphology-based sinkhole detection methods, and implement these by developing tools that can be used in standard geographic information systems (GIS). The methodology used in this research allows for the development of a robust LiDAR-based existing sinkhole detection and mapping toolset that provides an adequate degree of accuracy and at the same time maximizing the ability to assist inspectors with varying expertise. As an additional benefit, the developed tools can also be used to detect and map playas and other surface depression hazards.

4. METHODOLOGY

This research project's methodology includes: (1) data acquisition and preparation; (2) preliminary sinkhole detection and map development; (3) final sinkhole detection and map development; and (4) sinkhole map accuracy validation.

4.1. Data Acquisition and Preparation

The project team focused its study on the geographical region of Southeast New Mexico (City of Roswell and City of Carlsbad area or known as Roswell-Carlsbad district) because this area typifies karst topography. Roswell-Carlsbad district is a large gypsum karst area. As an example, many potential massive sinkholes can occur in Carlsbad (Figure 1). Highways in this area have signs warning drivers to be wary of potential sinkholes (Figure 2). In addition, this research project's technical advisors, including New Mexico Department of Transportation (NMDOT) – District 2 and National Cave and Karst Research Institute (NCKRI), are also located in this area. Specially, NMDOT District 2 is located in the City of Roswell, while NCKRI is located in the City of Carlsbad. They have ground-surveyed or digitized sinkhole data and maps that were planned to be used as ground-truth data for sinkholes in this area. However, due to time constraints, the project team was only able to obtain digitized ground-truth data from NMDOT District 2.

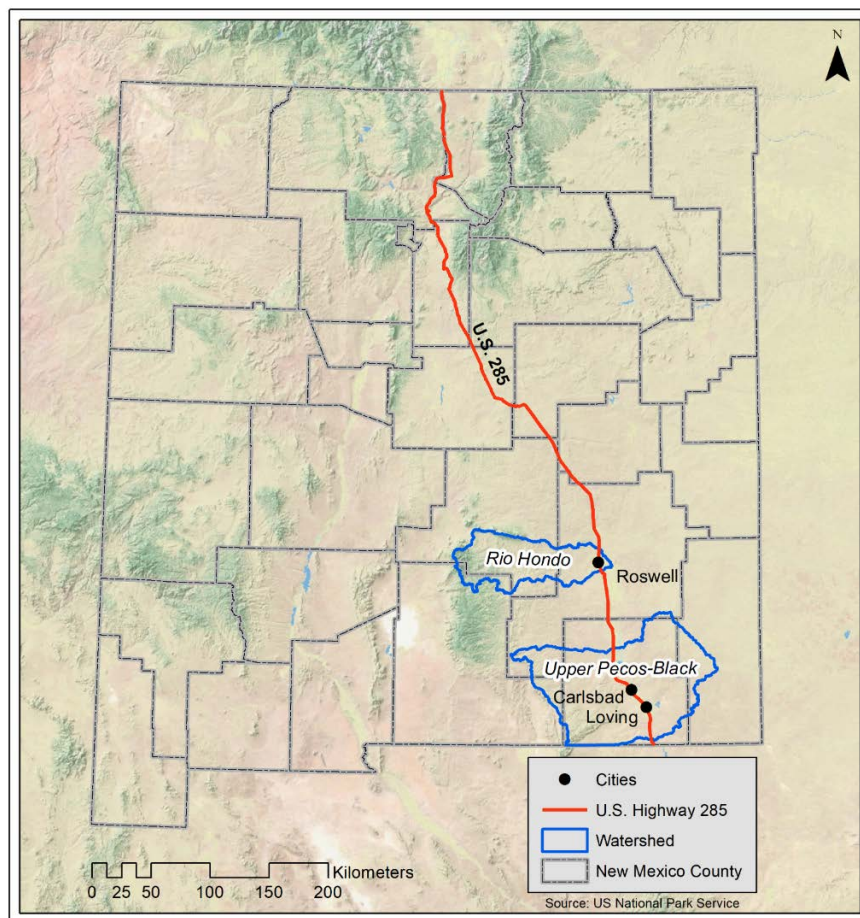


Figure 1. Roswell-Carlsbad District and Rio Hondo and Upper Pecos-Black Watershed, New Mexico.



Figure 2. Signs warn drivers to be wary of potential sinkholes near Carlsbad (adopted from Carlsbad Current Argus).

As part of the 3DEP, in 2014 the Federal Emergency Management Agency (FEMA) Region 6 collected Quality Level (QL2) airborne LiDAR data for the Rio Hondo Watershed (HUC08 sub-basin) located in Southeast New Mexico. Figure 1 shows the boundary of the Rio Hondo Watershed. This watershed includes most of the City of Roswell. It revealed that NMDOT did not collect QL2 airborne LiDAR for City of Carlsbad and Loving Village in 2016 within the Upper Pecos-Black Watershed. However, NMDOT collected aerial images for the City of Carlsbad and Loving Village in 2016. The coverage of these aerial images are shown in Figure 3. Therefore, the project team used the FEMA collected QL2 airborne LiDAR for this research.

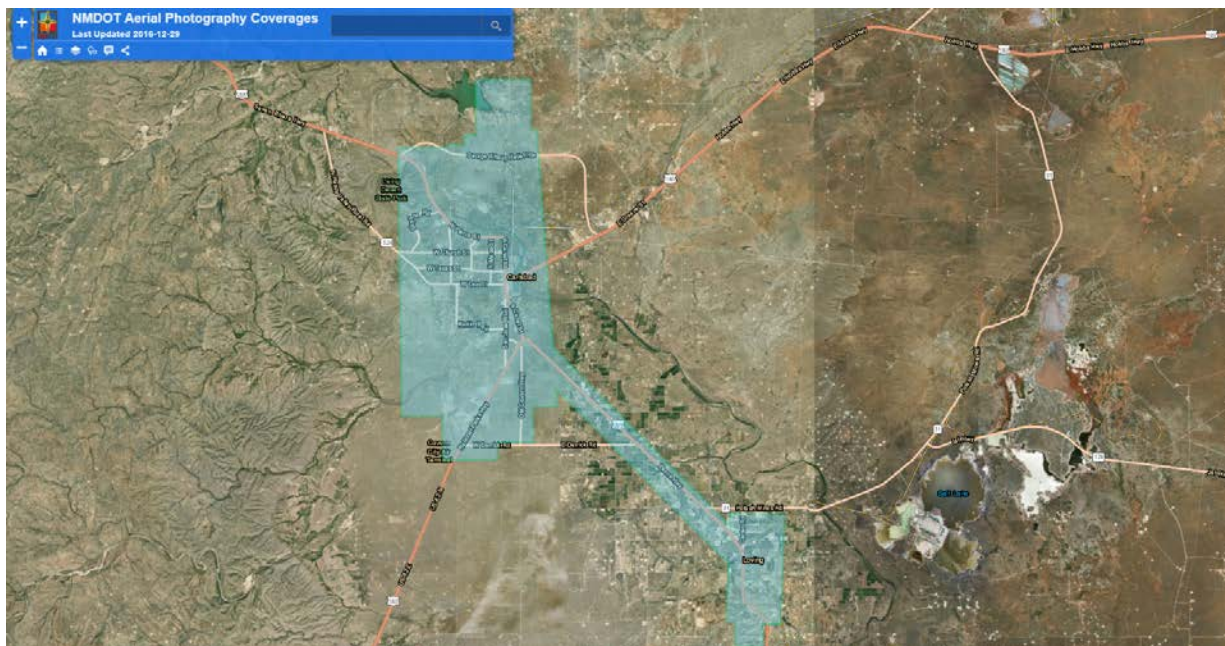


Figure 3. Aerial photo coverage for Carlsbad and Loving Village (adopted from NMDOT).

The airborne LiDAR for this area are in raw format – LAS format, files that have .las or .laz extension (compressed version of las). The .las files are typically very large, contain hundreds of

thousands points, known as a point cloud (Figure 4), and therefore, most times they are compressed to .laz files to reduce the file size by approximately 50%. LAS format is a binary format which can maintain information specific to the nature of the airborne LiDAR data while not being overly complex. The approximately 2 terabytes (TB) of airborne LiDAR data were divided into 2,411 separate las tiles of 1,500 m by 1,500 m at the maximum (or up to 225 hectares) in order to assist with data handling and processing. Files were projected into the UTM projection, Zone 13N, using the NAD83 datum and GRS80 spheroid. All elevation data were delivered in meters (Figure 5).

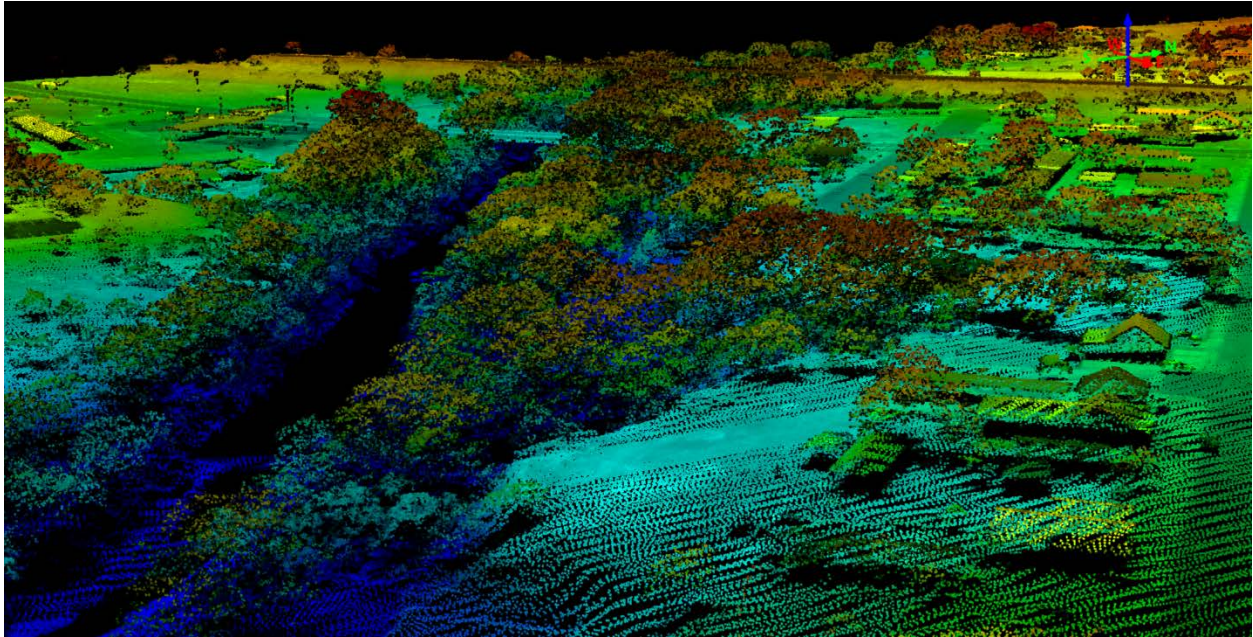


Figure 4. The display of a point cloud that covers a small portion of the City of Albuquerque.

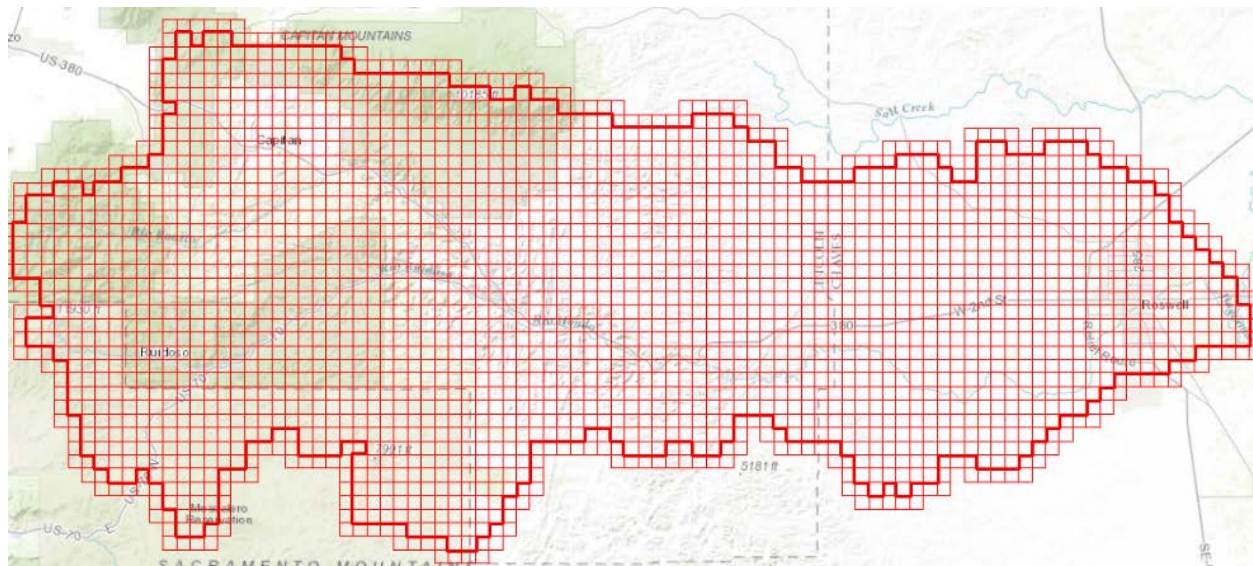


Figure 5. The Rio Hondo Watershed boundary (dark red) and individual LAS tiles (light red).

All individual LAS files were mosaicked into a more manageable LAS dataset since individual LAS file covers such a small portion of the ground. An LAS dataset is a standalone file that references one or more LAS files. In other words, a LAS dataset is a referenced container to a collection of LAS files. LAS datasets are designed to provide rapid access to large volumes of LAS files without the need for data conversion or importing. When working with a few LAS files that cover a small study area, LAS datasets may not be needed. However, when working with thousands of LAS files that cover an entire jurisdiction area, LAS datasets are the most effective solution because users can rapidly access to any specific LAS tile without looking for its spatial coverage information. LAS datasets are a proprietary format developed by ArcGIS – a widely adopted GIS software, which was used as the primary software for this research.

An LAS dataset can also have feature classes for surface constraints, such as breaklines, water polygons, and area boundaries. Up to three files can be associated with an LAS dataset, including LAS data file (.lasd), LAS auxiliary file (.lasx), and LAS projection file (.prj). The .lasd file only stores references to actual LAS files and surface constraints. The .lasx file provides a spatial index structure that helps improve the performance of an LAS dataset. If LAS files do not have a spatial reference or have an incorrect spatial reference defined in their header files, a .prj file can be created for the LAS dataset. In that case, the new coordinate system information in the .prj file will take precedence over the spatial reference in the header section of the LAS files.

Once the LAS dataset is created, it can be used to create digital elevation models (DEMs) and digital surface models (DSMs). A DEM, also known as digital terrain model (DTM), is a digital representation of bare ground surface terrain, while DSM is a digital representation of the mean sea level (MSL) elevations of the reflective surfaces of trees, buildings, and other features elevated above the bare earth. Figure 6 shows the difference between DTM and DSM.

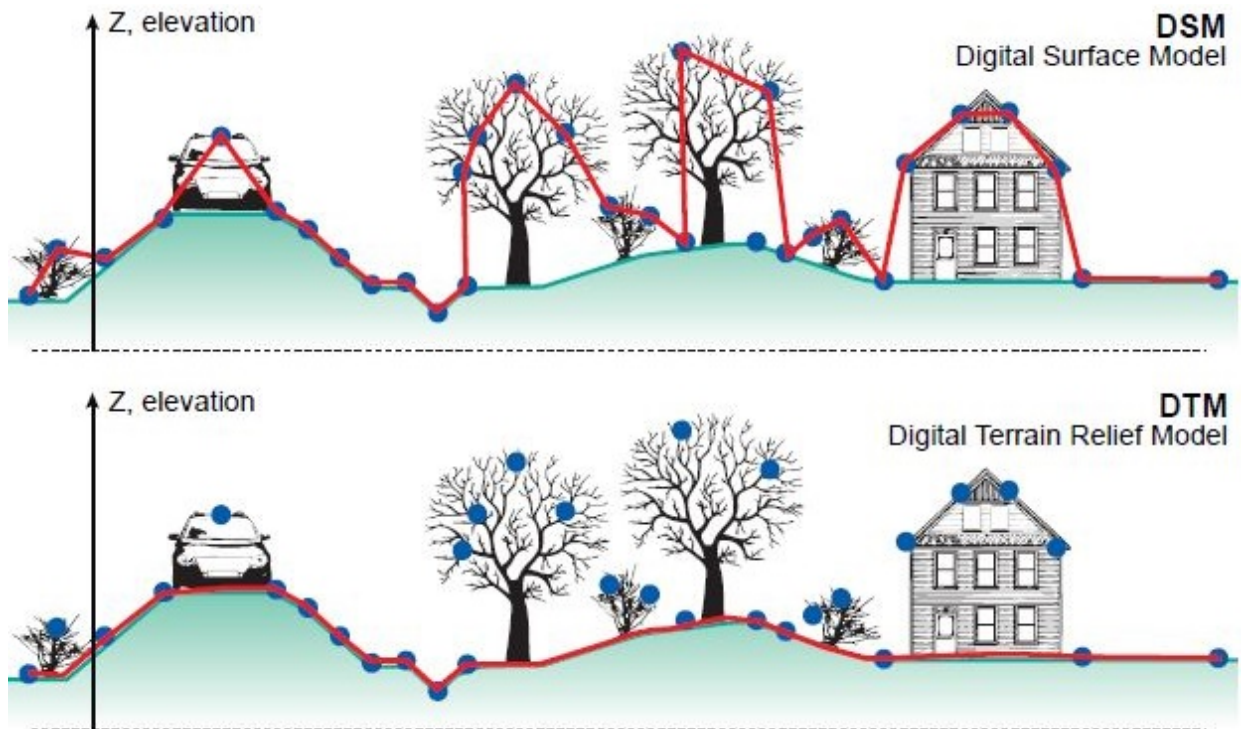


Figure 6. An illustration of the difference between DTM and DSM (adopted from 3DMetrica).

To enable the detection of sinkholes, DEMs (or DTM) are necessary because sinkholes depressions in the Earth's surface. Because the Rio Hondo Watershed is a very large area, the project team processed 40 LAS tiles that are in close proximity to the East side of the City of Roswell and the size of the study area is approximately 90 km² (Figure 7).

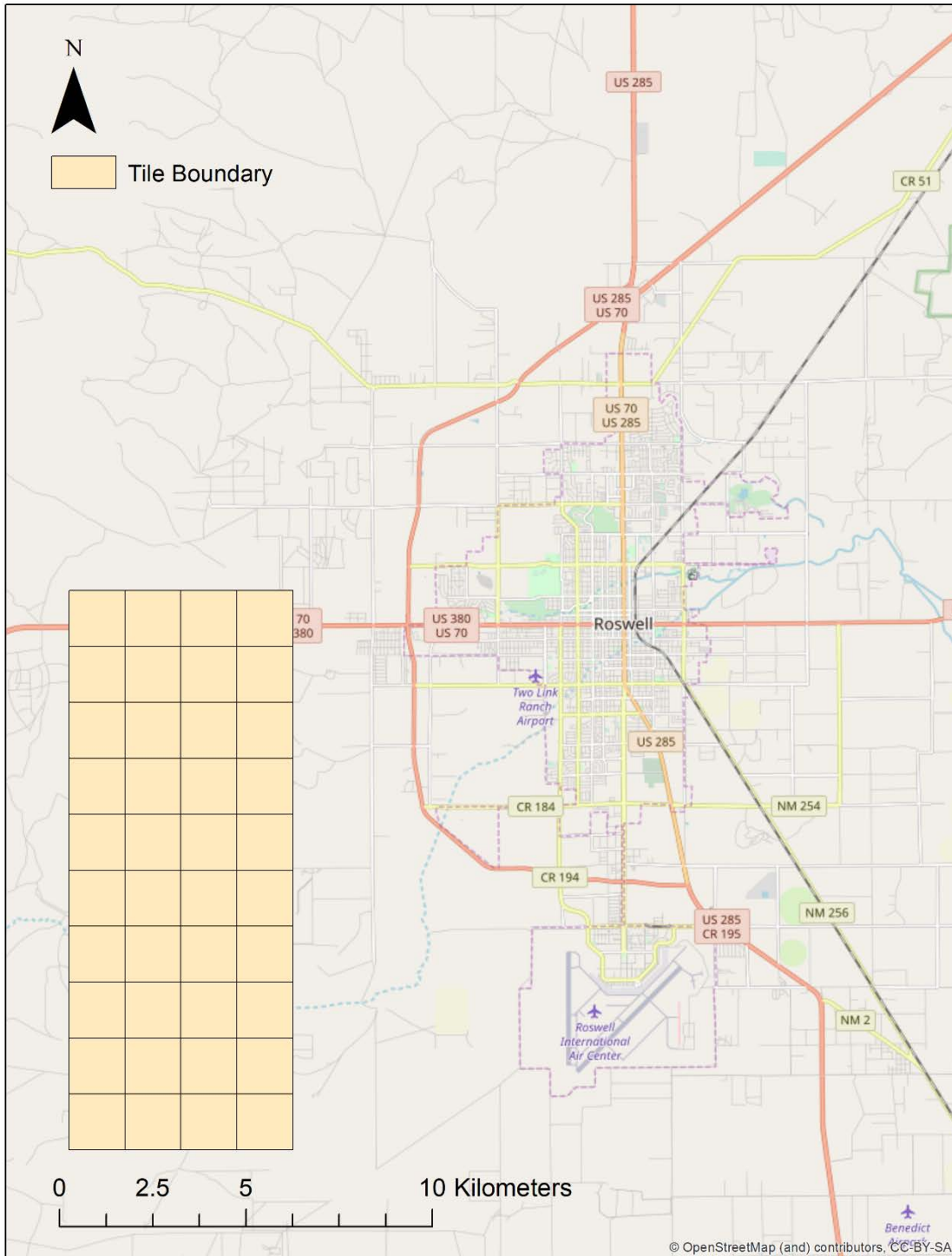


Figure 7. The processed DEMs tile for the study area; 40 tiles with a total area of 90 km².

4.2. Preliminary Sinkhole Detection and Map Development

Geometrically, sinkholes are depressions in the Earth's surface surrounded by terrain (12). Their relative subsidence from the Earth's surface level characteristic leads the sinkhole detection principle focuses on identifying the distinct shape-transition boundary in reference to the surrounding terrain (Figure 8). For this step, the project team focused on detecting all the possible locations for existing sinkholes and presenting them on a map.



Figure 8. Sinkhole detection principle of identifying the distinct shape-transition boundary in reference to the surrounding terrain (adopted from NCKRI).

In GIS, surface depressions are referred to as sinks, which are a cell of a set of spatially connected cells whose flow direction cannot be assigned any of the eight valid values for a flow direction (Figure 9), which can occur when all neighboring cells have higher values than the processing cell or cells in the middle. The project team used this concept to identify all possible locations of surface depressions, which is described in the following.

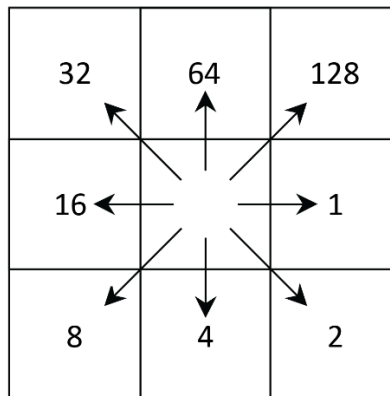


Figure 9. Eight flow directions for all neighboring cells of a focal cell.

The project team used pixel-based feature extraction technique to identify all possible locations of surface depressions. First, all sinks were identified. It should be noted that sinks are not sinkholes. In GIS, a sink is defined as a cell with an undefined drainage direction – no cells surrounding it are lower than it. Second, all sinks were filled (Figure 10). Third, the original DEM was subtracted from the filled DEM to detect the location and depth of all possible sinkholes to generate the preliminary sinkhole map. A geoprocessing tool (Figure 11) compatible with ArcGIS (Figure 12) was developed to automate the entire process mentioned above.

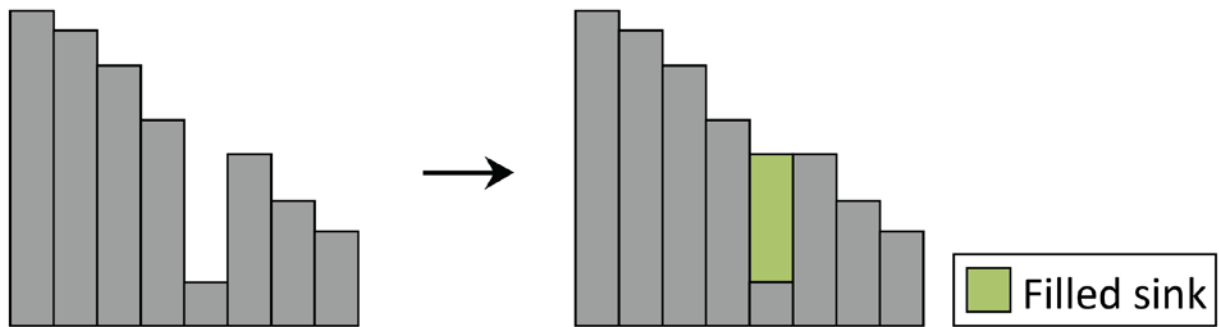


Figure 10. An illustration of the process of filling sinks.

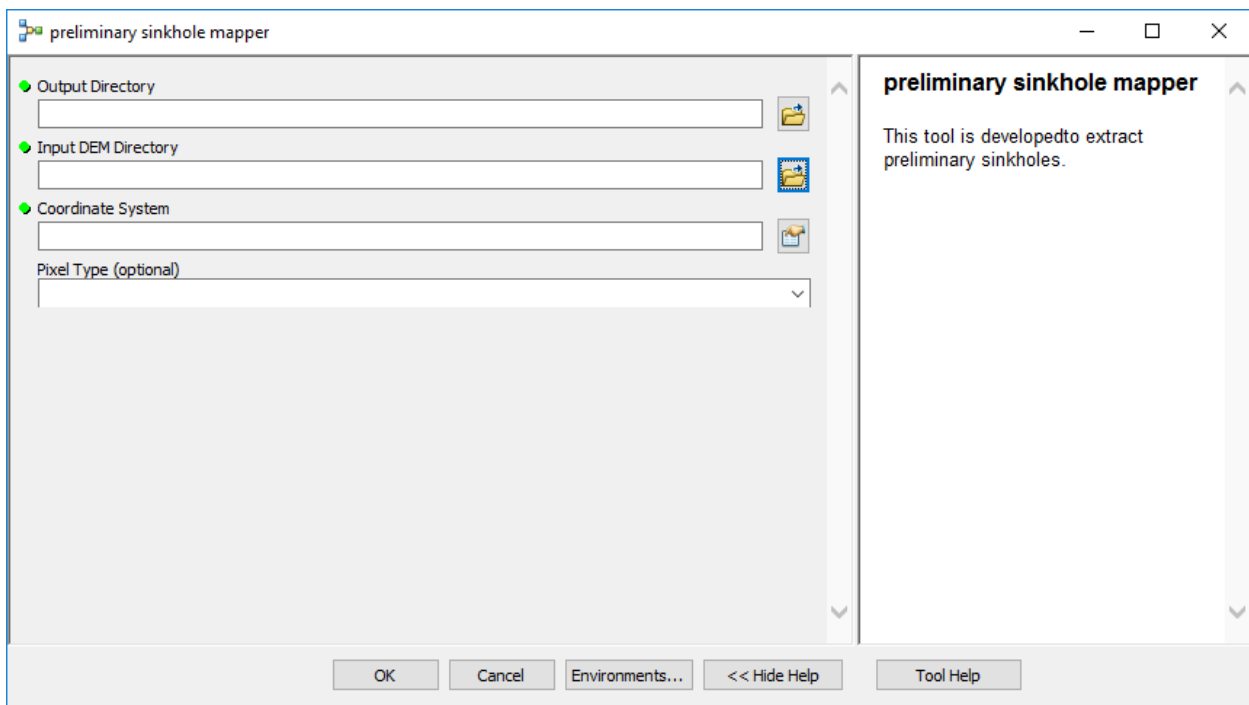


Figure 11. The Interface of the geoprocessing tool for preliminary sinkhole detection.

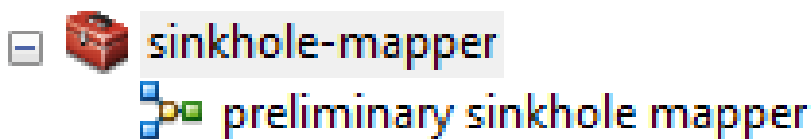


Figure 12. The toolbox that can be added to the ArcGIS software interface.

4.3. Final Sinkhole Detection and Map Development

For this task, the project team focused on improving the preliminary sinkhole map through two methods. The first method was using context-based feature extraction (e.g., OBIA) to remove false positively detected sinkholes. The second method was using morphology-based feature extraction (e.g., size) to remove false positively detected sinkholes. Detailed discussion of these two methods are provided in the following section.

4.3.1. Context-based Feature Extraction

Prior to discussing context-based feature extraction, it is necessary to discuss remote sensing image classification. The classification of remotely sensed imagery (e.g., aerial photos, satellite imagery, and LiDAR derived DEM and DSM) is the process of assigning land use and land cover (LULC) classes to pixels (Figure 13). Generally, remote sensing image classification can be divided into two general approaches, including pixel-based classification and object-based classification. Although pixel-based analysis has long been the mainstay approach for classifying remote sensing images, OBIA has become increasingly popular over the last decade.



Figure 13. An illustration of remote sensing imagery classification (adopted from Land Info).

For pixel-based classification, individual image pixels or cells are analyzed by the spectral information (spectral values) they contain. That being said, pixel-based classification methods explore the spectral information for assigning a pixel to a class according to the spectral difference between the classes. There are two types of pixel-based classification methods, including unsupervised classification and supervised classification. In unsupervised classification, the classes are created based on spectral analysis of pixels without the user providing training sites (group of sample pixels). Then the image processing software will use established algorithms (e.g.,

k-means and ISODATA) to determine which pixels are related and group them into classes. The users can specify which algorithm to use and the desired number of output classes.

Conversely, supervised classification methods are based on the idea that a user can select training sites in an image that are representative of specific classes and then direct the image processing software to use these training sites as references for the classification of the remaining pixels in the image through the use of various algorithms (e.g., minimum distance to means and maximum likelihood). Training sites are selected based on the knowledge of the users. The users can also set the bounds for how similar pixels have to be to group them. These bounds are often determined based on the spectral characteristics of the training sites with a certain increment or decrement. The users can also determine the number of classes by creating matching training sites.

Many remote sensing analysts use a combination of supervised and unsupervised classification processes to develop output classes. However, these approaches have a common drawback – they only use spectral information; that said, they do not exploit spatial and contextual information. For example, when using pixel-based classification methods for an image, a dry stock pond will be false positively detected as a sinkhole due to its corresponding pixels’ spectral similarity. Therefore, oftentimes users will receive a classified images with “salt and pepper” issues.

Unlike pixel-based classification, OBIA operates on objects consisting of many homogenous pixels grouped together in meaningful way through segmentation. In addition to the spectral information used in pixel-based classification methods, OBIA exploits shape and spatial proximity needed for classification. One of the most significant advantages of segmentation is that it creates objects representing land cover types that may be spectrally variable at the pixel level and thus eliminates the “salt and pepper” issues associated with per-pixel classification (Figure 14). Another advantage of OBIA is its ability to leverage feature shapes, topological features, and hierarchical structures to group pixels to objects.



Figure 14. Pixel-based image analysis classification vs. OBIA classification.

One disadvantage of OBIA is that the segmentation process and subsequent calculation of the topological relationships between objects can use a large amount of computer resources. However, recent advances in computing have provided adequate processing power to most computers. Therefore, this disadvantage of OBIA has limited impact on this research project.

Many algorithms exist for the segmentation process, and these algorithms can operate on many more object-related features than typically available with pixel-based approaches. These algorithms include Region Growing, Watershed Detection, and Mean Shift. Region Growing focuses on finding similar pixels from a seed and neighboring pixels. Watershed Detection is mostly used for gray-scale images, and it treats images like a topographic surface to detect homogeneous pixels to group them as a watershed. Mean Shift is a local homogenization technique that concentrates on damping shading and tonality difference in localized objects to find the clustering of objects. Each algorithm has its strengths and limitations and there is no unanimous consent on the best one. The project team decided to use Mean Shift because it is compatible with ArcGIS which enables the automation of the entire process. The general process of Mean Shift algorithm is: (1) identifying regions with high density in feature space; (2) determining the centroid of each region; (3) moving the region to the location of the new centroid; and (4) repeat until convergence. Figure 15 shows the aforementioned Mean Shift segmentation process.

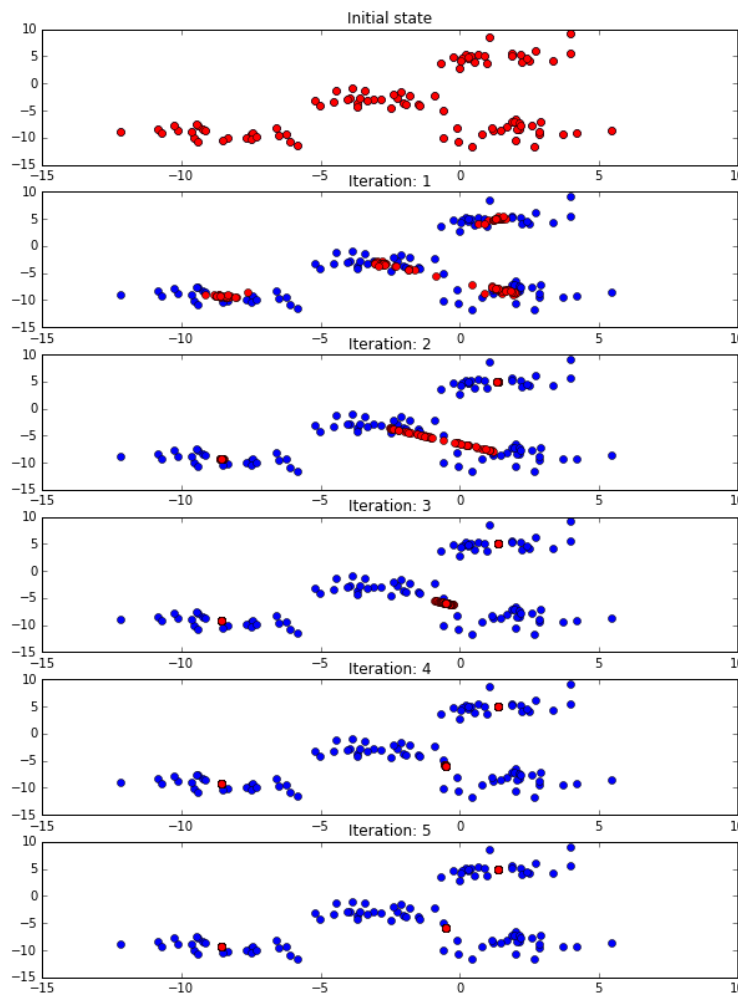


Figure 15. An illustration of the Mean Shift segmentation (adopted from chioka.in).

The detailed parameters for Mean Shift algorithm can be found Appendix B. With the aid of the Mean Shift segmentation, it is possible to delineate sinkhole boundaries more accurately. Being able to accurately delineate sinkhole boundaries is important to undertake many sinkhole inspection and decision-making actions. First, it enables a better estimation of a sinkhole’s basic morphometric characteristics, including, but not limited to, width, length, area, perimeter, depth, volume, elongatedness (length/width), and standard deviation of elevation. Second, it enables the characterization and investigation of each individual sinkhole.

In addition, other context information, including, but not limited to, soil types, infrastructure and political boundary, vegetation, LULC, and hydrology, can be used to further improve sinkhole detection accuracy. Table 1 lists all the context information and how to appropriately use them. Users should identify and obtain all available layers and composite them into a single layer and use it as a filtering layer to constrain the analysis to those areas where sinkholes are expected.

Table 1. Context information elements and matching application.

Context Data	Example Sources	General Applications
Shape and Spatial Proximity	LiDAR derived DEM	Delineate sinkhole boundary
Soil Types	USGS geological maps	Sinkholes will not occur in certain types of soil such as gypsum-rich soils
Infrastructure Boundary	Infrastructure footprint boundary maps	Remove objects that false positively detected as sinkholes such as stadiums
Vegetation	USDA National Agricultural Imagery Program (NAIP)	Vegetation located on top of sinkholes appears circular shape
LULC	USGS National Land Cover Database (NLCD) maps	Remove objects that false positively detected as sinkholes such as dry ponds
Hydrology	USGS National Hydrology Database (NHD)	Remove objects that false positively detected as sinkholes such as dry ponds

4.3.2. Morphology-based Feature Extraction

The project team also used morphology-based feature extraction concept to further improve the accuracy of sinkhole detection. The premise is that sinkholes will have a specific range of surface area for a certain area. For example, if sinkholes are expected to have a surface area ranges from 100 m² to 10,000 m², users should be able to leverage this area-range to filter the sinkhole detection results. That being said, when the area-range filter is applied, any sinkholes that are smaller than 100 m² and larger than 10,000 m² will be identified, selected, and discarded. The project team worked with NCKRI and it revealed that the aforementioned size range is operationally effective for the project’s study area. At the time of the report writing, the project team decided to use only sinkhole size in terms of area to conduct the morphology-based feature extraction. Other morphometric characteristics, including width, length, perimeter, depth, volume, elongatedness (length/width), and standard deviation of elevation, have not been widely used for sinkhole extraction yet due to limited understanding of their appropriate usage. That being said, a sinkhole could have any width, length, perimeter, depth, volume, and possibly elongatedness.

4.3.3. Methodology Workflow Summary

This section summarizes the workflow of the aforementioned steps. A workflow chart can assist in streamlining and automating repeatable tasks, minimizing the chances for generating errors, and maximizing overall sinkhole detection and mapping efficiency. A workflow can also help the audience understand the complete process of the methodology. Figure 16 shows the flowchart of the methodology. It should be noted that the flowchart includes data preparation, preliminary sinkhole detection and mapping, and final sinkhole detection and mapping.

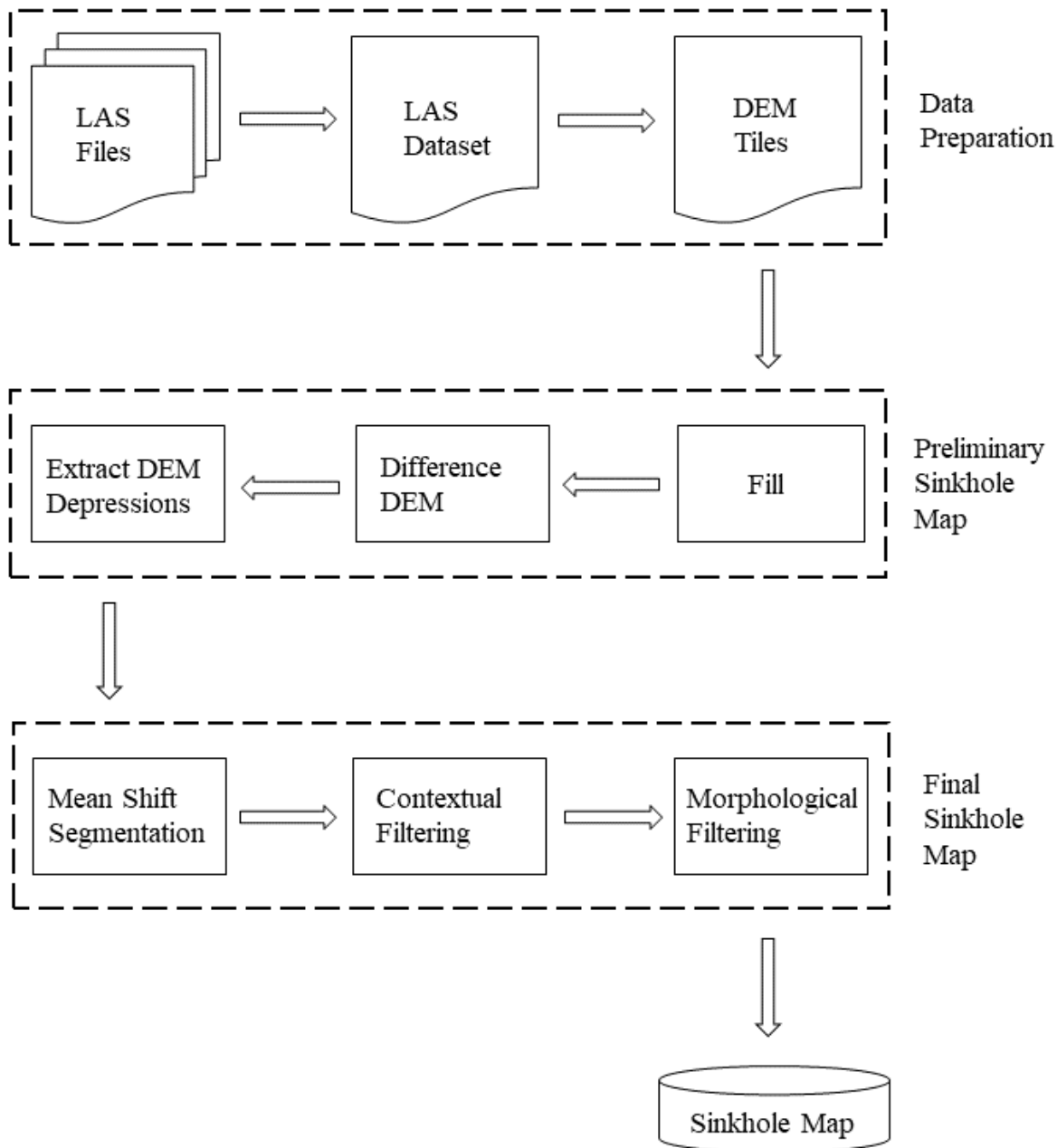


Figure 16. Flowchart of the methodology.

Two tools were developed to accomplish the aforementioned process. The first tool is the “DEM Creator” and it is designed to create DEM tiles if users do not have them readily available. The second tool is the “Sinkhole Extractor” and it is designed to extract sinkholes and map them. Figure 17 shows the Sinkhole Toolbox and the two aforementioned tools reside in. Detailed user interface for the two aforementioned tools will be provided in Section 5 Analysis and Findings.

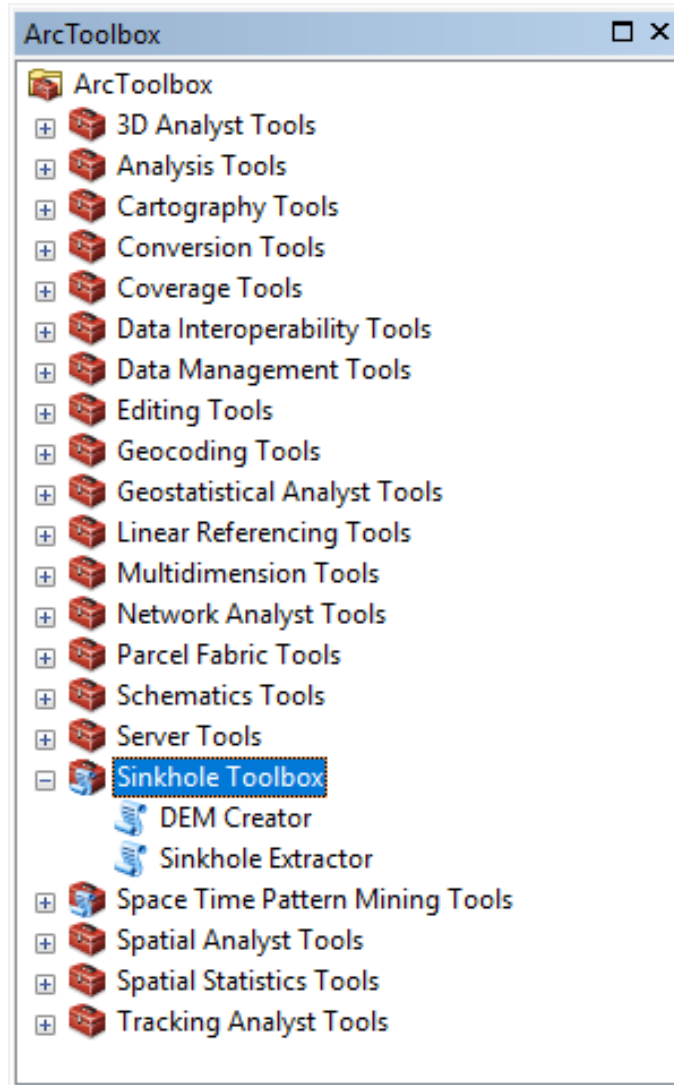


Figure 17. The Sinkhole Toolbox that is compatible with ArcGIS.

4.4. Sinkhole Map Validation

This step focused on validating the accuracy of sinkhole detection and mapping. Sinkhole data collected through aerial photo digitization were used as the ground-truth data and then compared them to the sinkhole map generated from airborne LiDAR data analysis. Cohen’s Kappa statistics was used to measure the overall agreement between the LiDAR detected sinkholes points and a set of ground-truth sinkhole points (aerial photo digitized sinkholes). At individual sinkhole level, the project team planned to use either Paired Student T-Test (if the data are parametric) or Wilcoxon Signed Rank Test (if the data are non-parametric) to examine if LiDAR detected sinkholes and ground surveyed sinkholes have statistically similar morphometric measurements.

5. ANALYSIS AND FINDINGS

5.1. Data Preparation Results

As mentioned in Section 4.1., the LAS tiles were processed to generate a LAS dataset, which was then used to create DEM tiles. It should be noted that many LAS tiles can be used to create a single mosaicked LAS dataset or each LAS tile can be used to create a single LAS dataset. In the latter case, the amount of LAS datasets will be equal to the amount of LAS tiles. This project created a single LAS dataset for all LAS tiles, and then created a DEM from the LAS dataset. Operationally, both methods will work. However, it is suggested that creating a single LAS dataset for many LAS tiles. This is because each individual LAS file covers a small portion of the ground and mosaicking them into a single and more manageable LAS dataset can lead to more effective data processing.

Considering that oftentimes users will have LAS files instead of DEMs, the project team developed an ArcGIS compatible tool to assist with data processing. This tool should be used if the users do not have existing DEM tiles that can be used to detect and map sinkholes. This tool has the following user interface (Figure 18), and it can be used to create DEM tiles from the LiDAR LAS tiles. As the figure shows, users can decide whether to input a single LAS tile or a folder of LAS tiles through the LAS Input Directory. Each individual LAS tile should be placed into a separate folder if the users decide to input a single LAS tile. Detailed instruction on how to use this tool is provided in Appendix B. It is suggested that the users adopt the default settings, and the few parameters that need users' input are the LAS Input Directory, Output Raster Directory, and the Output Coordinate System. The output DEM raster will be a single tile if the users input a single LAS tile. The output DEM raster will be also be single tile if the users input many LAS tiles, but the single DEM tile's spatial coverage will be equal to all input LAS tiles'.

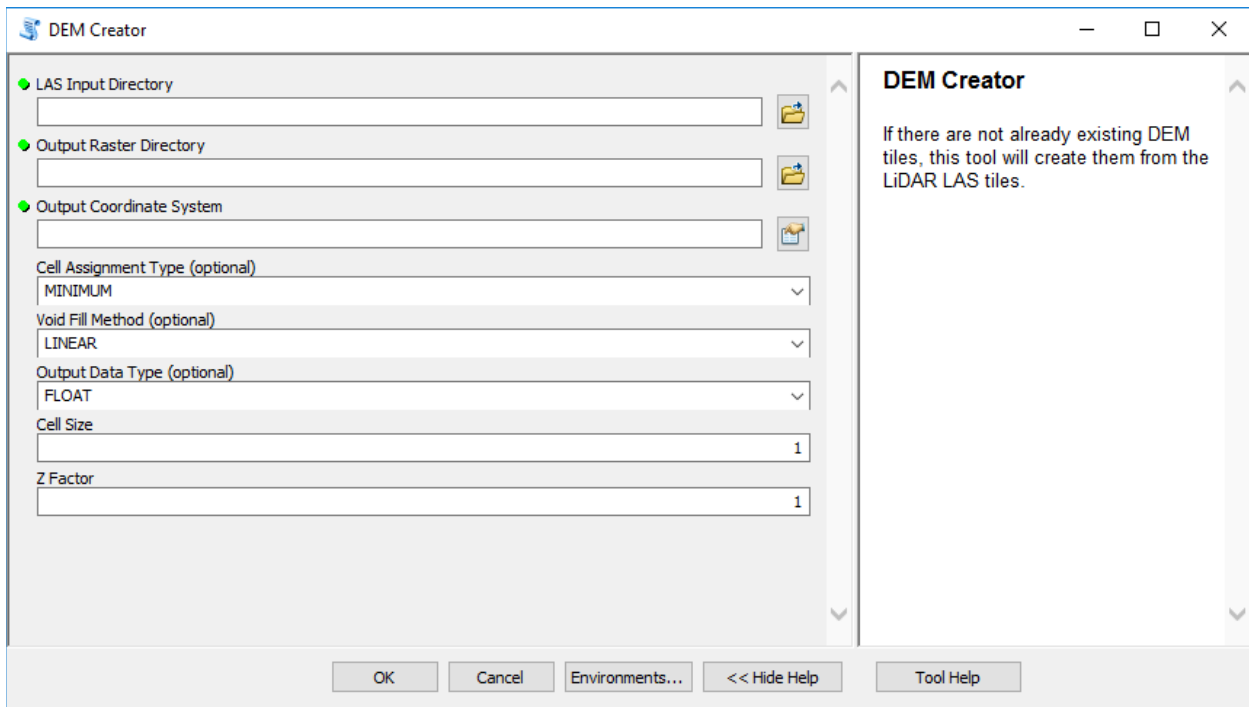


Figure 18. The user interface of the DEM Creator tool that is compatible with ArcGIS.

After running the DEM Creator tool, users will obtain a DEM raster layer as below. It should be noted that a single mosaicked DEM layer will be created even if users decide to use multiple LAS tiles as input for the DEM Creator tool (Figure 19). Individual DEM layers will be created separately if the users decide to use a single LAS tile as the input. That being said, the tool needs to run many times in order to obtain the same DEM shown in Figure 20 if a single LAS tile is used.



Figure 19. The mosaicked DEM layer when using multiple LAS tiles as input.

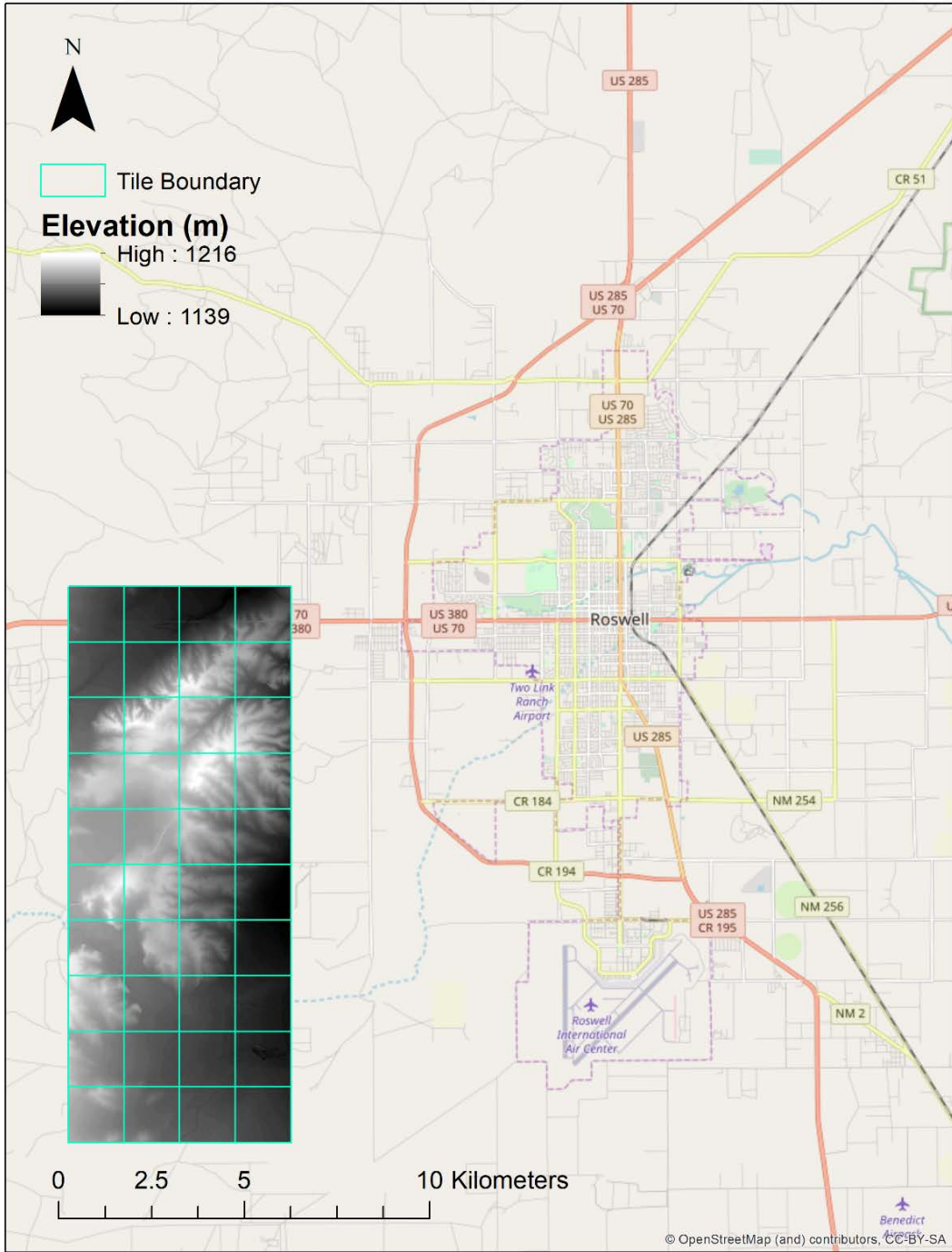


Figure 20. The individual DEM layers when using a single LAS tile as input.

5.2. Preliminary Sinkhole Detection and Mapping Results

As mentioned in Section 4.2., when using the preliminary sinkhole detection and mapping tool, users need to provide the input DEM tile or tiles and select the output location to save the preliminarily detected sinkhole layer. It should be noted that users can use either a mosaicked DEM tile or individual DEM tiles, depending on the availability of the data. Figure 21 shows the

output sinkhole layer from this tool. As it shows, the detected sinkholes are distributed across the terrain. That being said, the detected sinkholes appear all over the map. However, in reality the density of sinkholes will be less than this. The next task, final sinkhole detection and mapping, was focused on improving the accuracy of this preliminary sinkhole map.

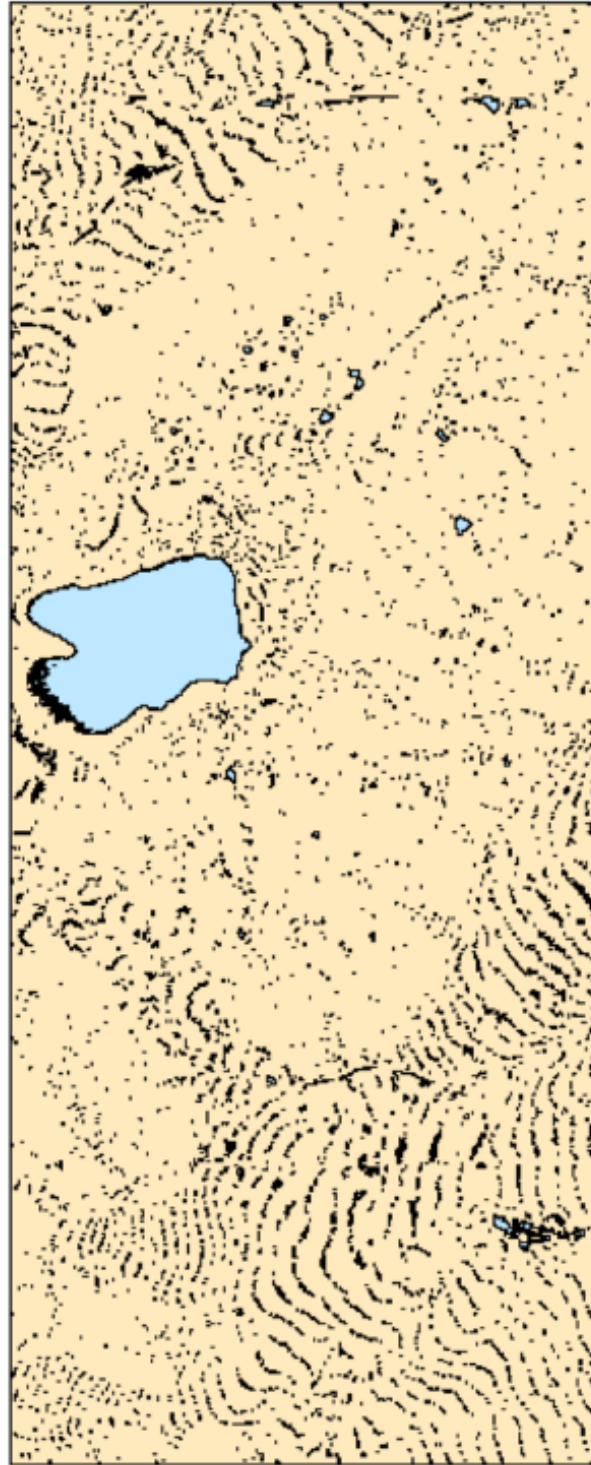


Figure 21. The output of the preliminary sinkhole detection tool.

5.3. Final Sinkhole Detection and Mapping Results

As mentioned in Section 4.3., the project team developed a final sinkhole detection and mapping tool based on context-based feature extraction techniques and morphology-based feature extraction techniques. Figure 22 exhibits the user interface of this tool. As the figure shows, users can choose the output path through the Output Directory parameter, select the input DEM tiles through the DEM Input Directory parameter, and choose the output sinkhole polygon file's coordinate through the Output Coordinate System parameter. The Extract Mask parameter is optional, but it can be used to constrain the analysis to those areas where sinkholes are expected. The Extract Mask Parameter should be used when users know which portion of the study area have existing sinkholes. This will be an empirical knowledge developed based upon the familiarity of a certain area. As mentioned in Section 4.3.1, users should identify and obtain all available layers and composite them into a single layer and use it as a filtering layer to constrain the analysis to those areas where sinkholes are expected.

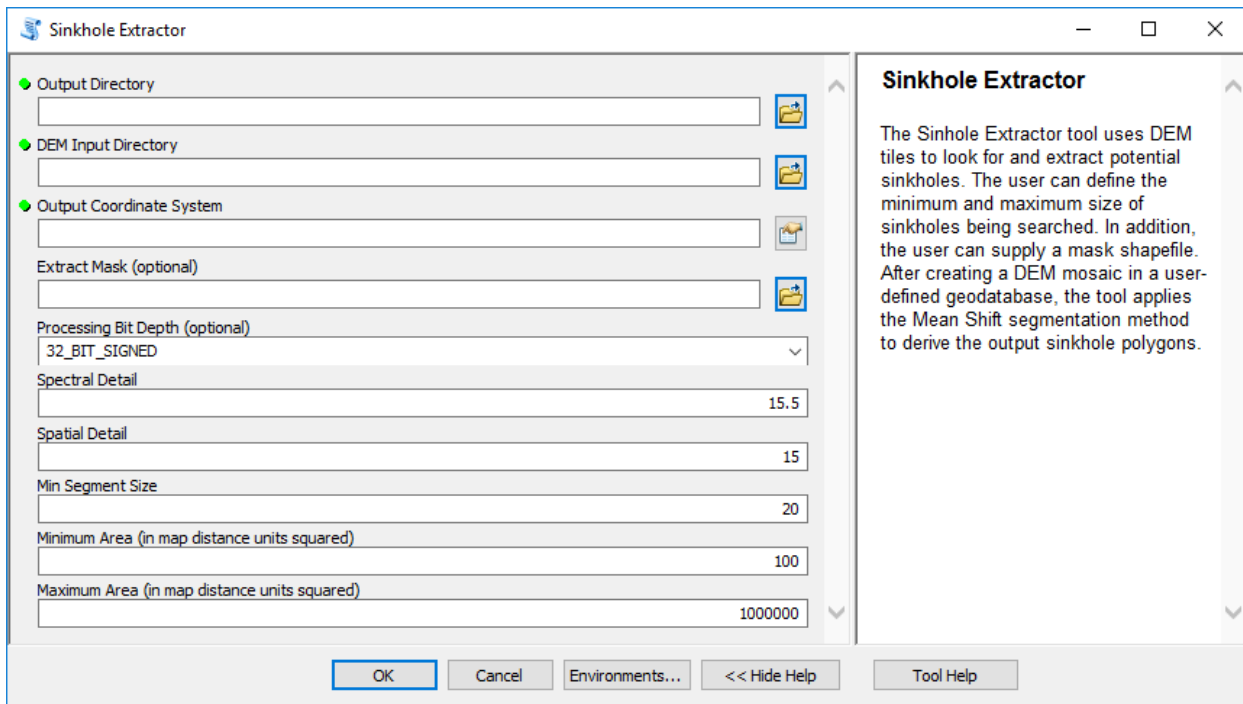


Figure 22. The user interface of the Sinkhole Extractor tool that is compatible with ArcGIS.

The Processing Bit Depth parameter is used to control the output bit depth of the DEM mosaic. Three bit depth values are available, including 32-bit Signed, 32-bit Float, and 64-bit. Their specific applications are summarized in Table 2. Users should use 32-Bit Signed, 32-Bit Float, and 64-Bit for a study area that is located in a region where the terrain is relatively flat, rolling, and steep, respectively. For this research, the study site is located in a relatively flat area, and therefore, 32-Bit Signed was used. More detailed information about this tool is provided in Appendix B.

Table 2. Bit depth options and their applications.

Bit Depth	Applications
32-bit Signed	When selecting 32-Bit Signed, the output DEM mosaic will truncate the elevation values to the nearest integer value which will output a smaller (file size) DEM mosaic file but will lose elevation details.
32-bit Float	When selecting 32-Bit Float, the output DEM mosaic will preserve the elevation values which will output a larger DEM mosaic but preserve the elevation detail.
64-bit	When selecting 64-Bit, the output DEM mosaic will also preserve the elevation values which will output a larger DEM mosaic but preserve the elevation detail. This should be selected if the 32-Bit Float option cannot provide enough value range.

Users can also specify the minimum and maximum sinkhole surface area to further filter the results. This is the minimum or maximum that will be found in the output sinkhole polygon file and measured in the square distance units in the map. This method is used as the morphology-based feature extraction. The default value for the minimum area and maximum area is 100 and 1,000,000 of the map unit squared (e.g., square meters). The project team worked with NCKRI and it revealed that the size range of 100 m² and 30,000 m² is operationally effective for the project's study area. Figure 23 shows the detection result when using 32-Bit Float and 100 m² and 5,000,000 m² limit.

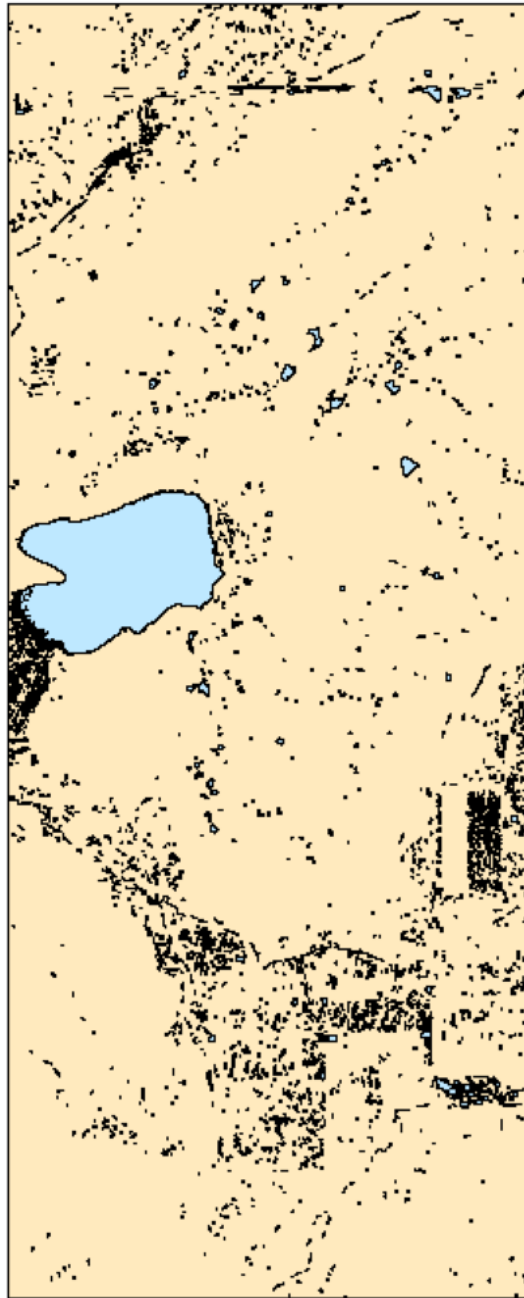


Figure 23. Sinkhole detection result when using 32-Bit Float with minimum 100 m² and maximum 5,000,000 m² area limit.

As Figure 23 shows, the detected sinkholes are still distributed across the terrain. That being said, the detected sinkholes still appear all over the map. This is because 32-Bit Float is not appropriate for relatively flat area. Figure 24 shows the detection result when using 32-Bit Signed and minimum 100 m² and maximum 5,000,000 m² area limit. As it shows, the result is more realistic. However, there is a very large sinkhole on the left side. Further exploration reveals that it is a reservoir instead of a sinkhole (Figure 25).

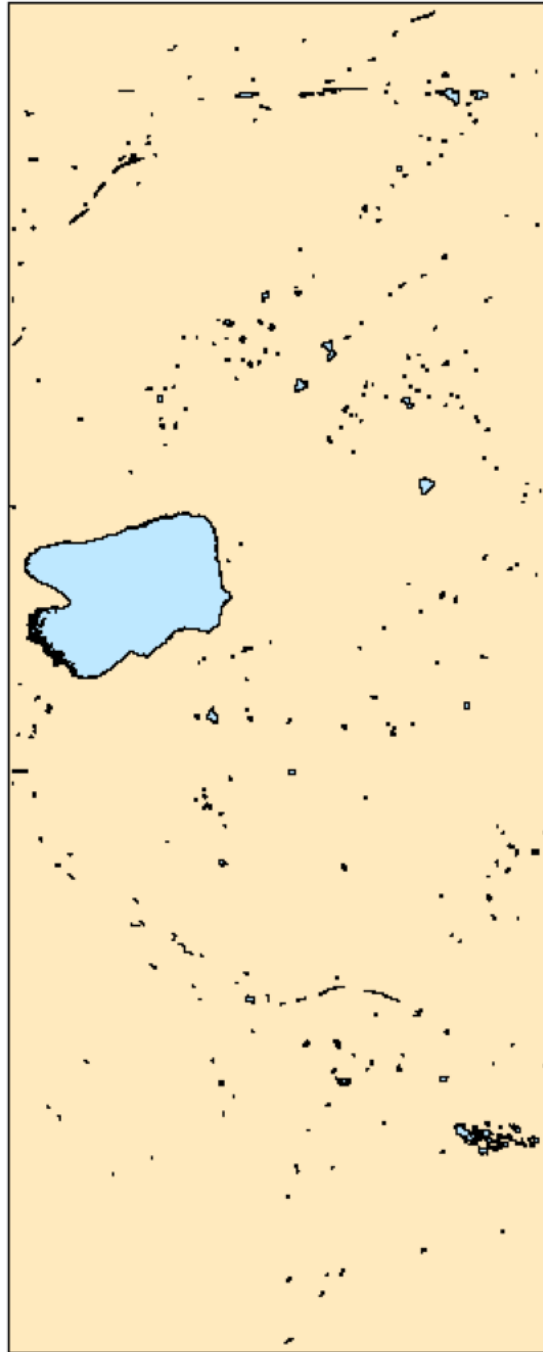


Figure 24. Sinkhole detection result when using 32-Bit Signed with minimum 100 m² and maximum 5,000,000 m² area limit.



Figure 25. A false positively detected sinkhole which is actually a reservoir.

This example shows that how important to use the morphology-based feature extraction technique to further filter the sinkhole detection results. Figure 26 shows the result when using the minimum and maximum area parameters to filter the sinkhole detection results. The project team used 100 m² for the minimum area and used 30,000 m² for the maximum area.

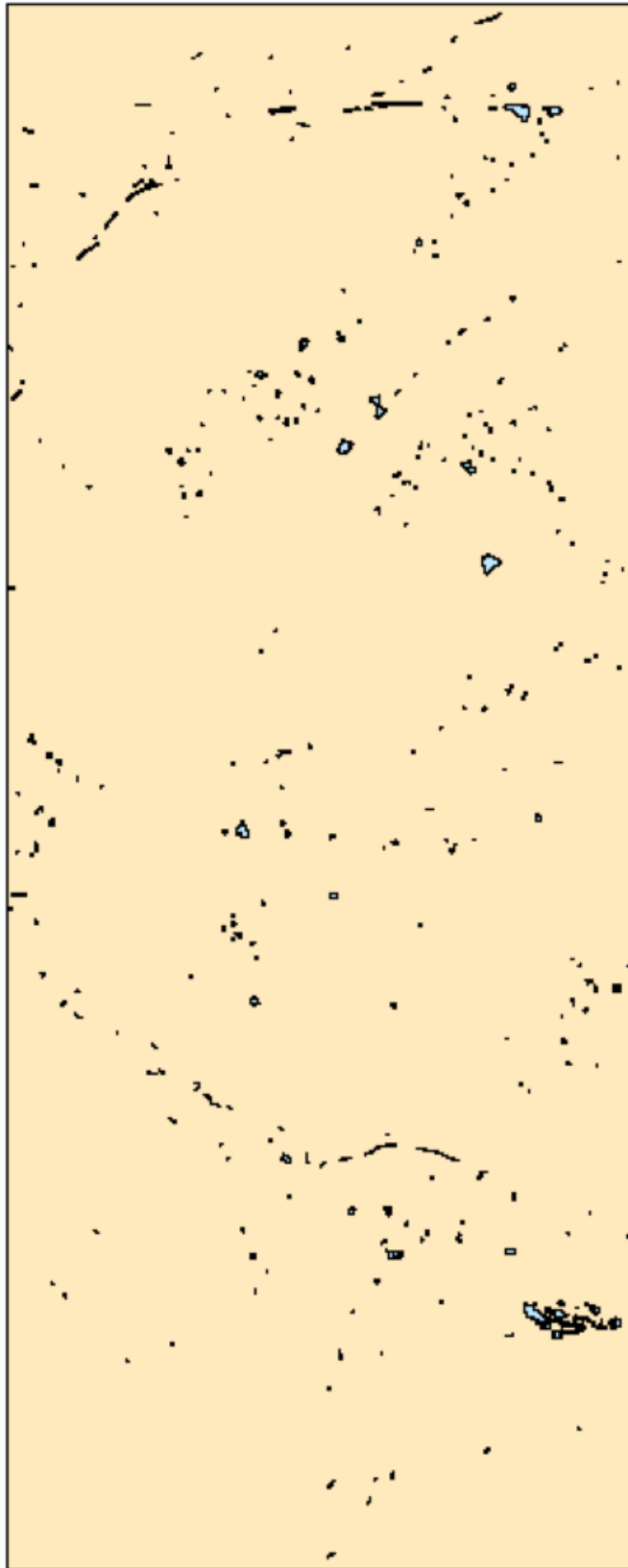


Figure 26. Sinkhole detection result when using 32-Bit Signed with minimum 100 m² and maximum 30,000 m² area limit.

As shown in Figure 25, there are 361 sinkholes that have been successfully detected and mapped. However, only certain areas within the study area are expected to have sinkholes. Another parameter that can be used to further filter the sinkhole detection result (with minimum 100 m² and maximum 30,000 m² area limit) is the Extract Mask. It can be used to constrain the sinkhole detection to those area where sinkhole are expected. Worked with NMDOT District 2, the project team was able to digitize a single shapefile layer and use it as a filtering layer to constrain the analysis to those areas where sinkholes are expected. Figure 27 shows the used extract mask.

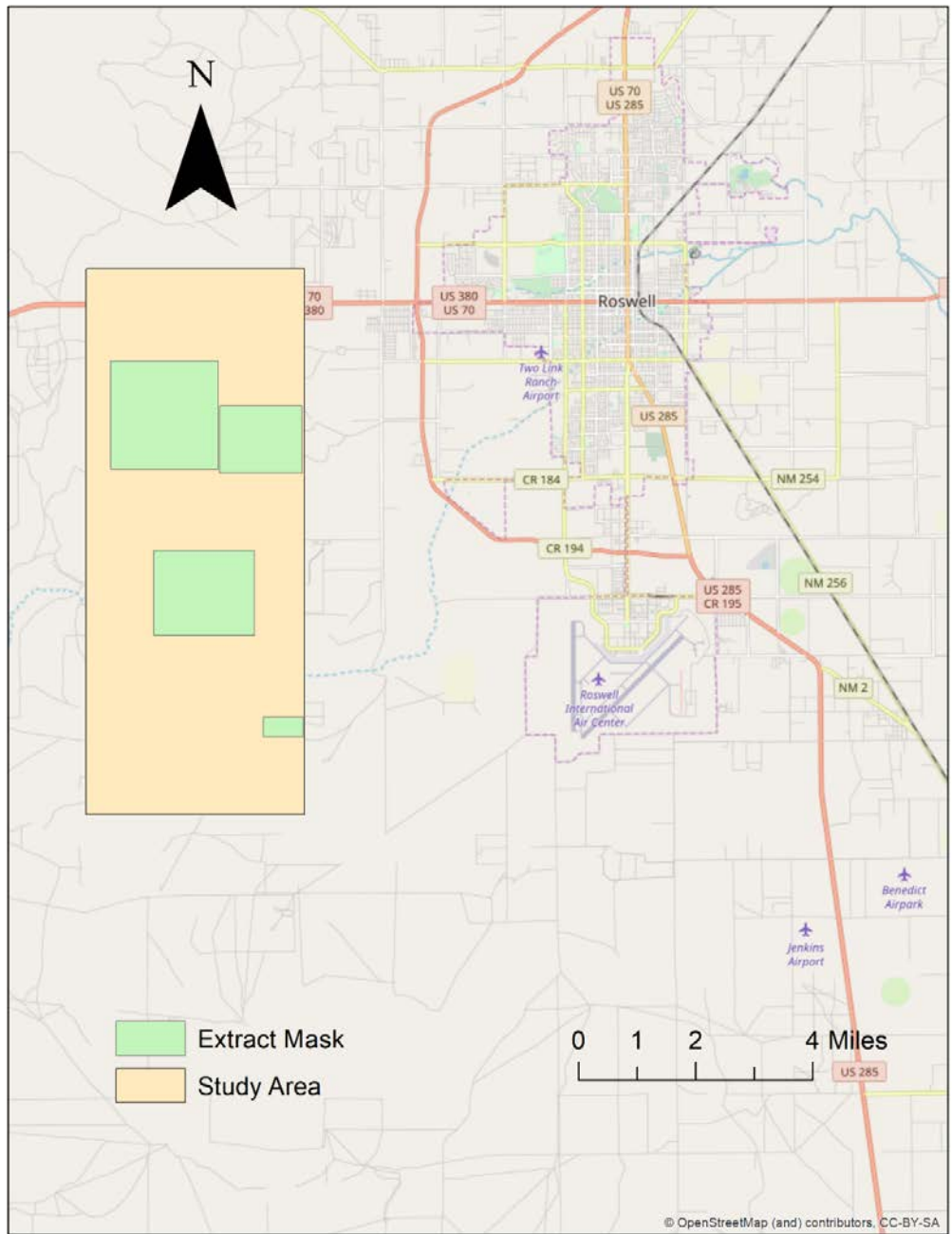


Figure 27. The location of the extract mask and the study area.

When the extract mask was applied, the sinkhole map is further improved. Figure 28 shows the improved sinkhole map when the extract mask was applied. There are 126 detected sinkholes in the extract mask zone. That being said, 235 sinkholes are further filtered out. Figure 29 shows the zoomed-in look of the area that is highlighted with red bordered box in Figure 28.

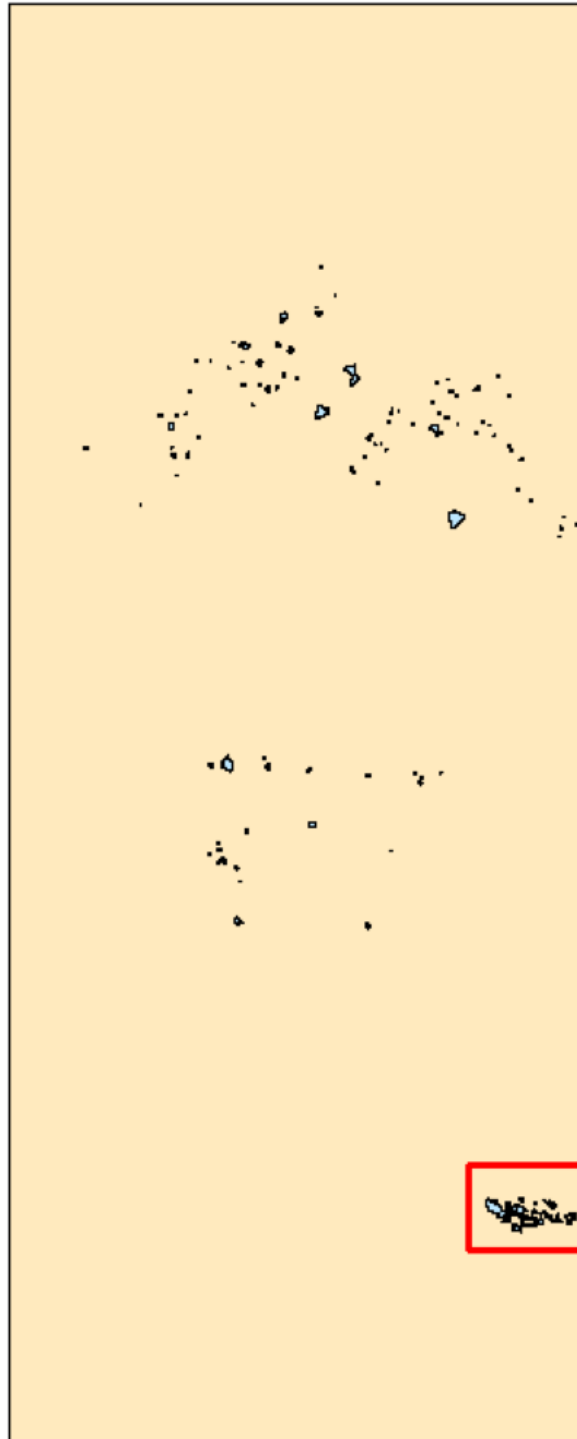


Figure 28. Sinkhole detection result when using 32-Bit Signed with minimum 100 m² and maximum 30,000 m² area limit and with the extract mask.

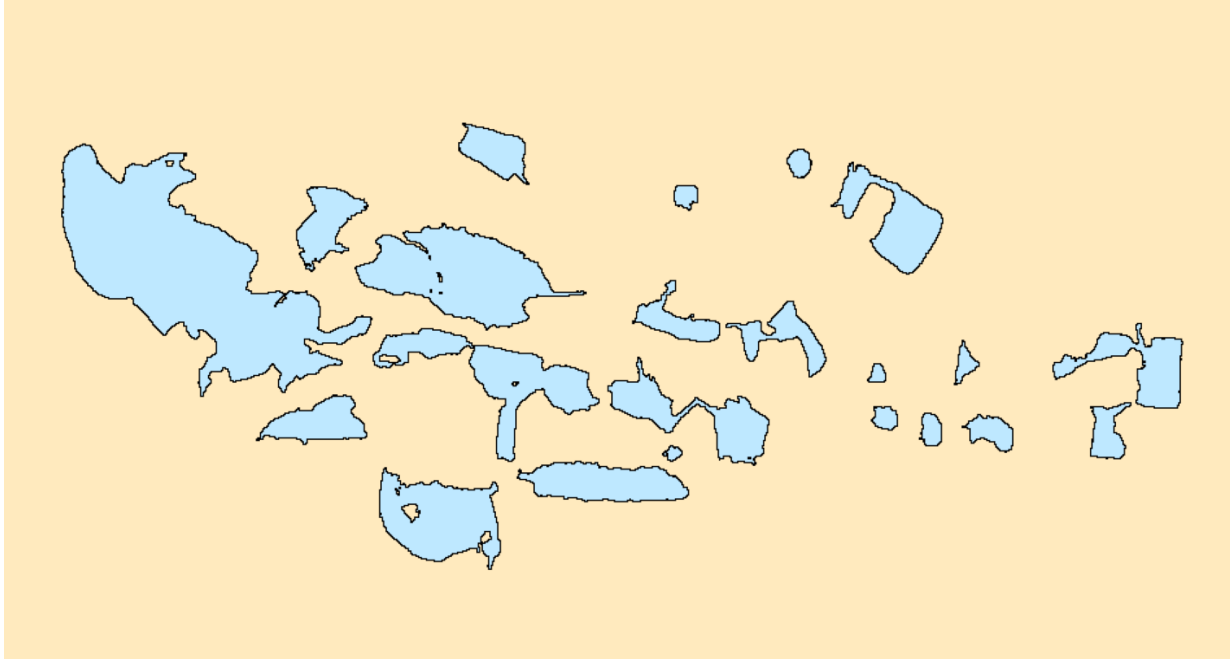


Figure 29. The zoomed-in look of the area that is highlighted in Figure 28.

5.4. Sinkhole Map Validation

This task was focused on validating the accuracy of detected sinkholes. The validation was conducted at two different levels, including the overall level and individual level. For the overall level validation, Cohen’s Kappa statistics was used to measure the overall agreement between the LiDAR detected sinkholes points and a set of ground-truth sinkhole points (aerial photo digitized sinkholes). For the individual level validation, Wilcoxon Signed Rank Test and Mann-Whitney U test were used to examine if LiDAR detected sinkholes and aerial photo digitized sinkholes have statistically similar morphometric measurements. Detailed discussion regarding these two validation processes is provided in the following section.

5.4.1. Overall Level Validation

Cohen’s Kappa is a statistics that can be used to measure inter-rater agreement for qualitative items. In order to perform the Cohen’s Kappa statistics, four parameters are needed, including True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). Their meaning in the context of this research is listed in Table 3.

Table 3. Cohen’s Kappa matrix.

		Ground-Truth Sinkholes	Ground-Truth Sinkholes
		YES	NO
LiDAR Detected Sinkholes	YES	TP	FP
LiDAR Detected Sinkholes	NO	FN	TN

The Spatial Join tool in ArcGIS was used to extract the TP value. Both the LiDAR detected sinkhole polygon layer and the ground-truth sinkhole polygon were uploaded to ArcGIS. There are 126 sinkholes in the LiDAR detected sinkhole polygon layer, while there are 121 sinkholes in the ground-truth sinkhole polygon layer. Subsequently, the Spatial Join tool was used to identify the amount of polygons that are common to both layers (i.e., intersected). The Spatial Join result revealed that 107 polygons were intersected, which indicated that the TP value is 107. This resulted in FP value being 19. This further resulted in FN value being 14 because the sum of TP and FN equals to the amount of sinkholes in the ground-truth sinkhole layer. As mentioned in Section 5.3, when using the extract mask, 235 sinkholes were further filtered out and they were not in both the LiDAR detected sinkhole layer and the ground-truth sinkhole layer. Therefore the TN value is 235. The project team also manually examined each of the 107 sinkholes is actually one-to-one matching. That said, the project team examined if a specific LiDAR detected sinkhole intersected with its matching ground-truth sinkhole but not any of the unmatching ones. The results revealed that the 107 sinkholes were all matched. The aforementioned results are summarized in Table 4.

Table 4. Cohen’s Kappa Matrix for LiDAR detected sinkholes.

		Ground-Truth Sinkholes	Ground-Truth Sinkholes	Total
		YES	NO	
LiDAR Detected Sinkholes	Yes	107	19	126
LiDAR Detected Sinkholes	No	14	235	249
Total		121	254	375

With the Cohen’s Kappa coefficient interpretation shown in Table 3, the project team conducted Cohen’s Kappa statistics calculation and the Cohen’s Kappa coefficient was 0.801. As a general rule of thumb, the coefficient value can be classified into five categories, as shown in Table 5 (23). As the table revealed, the Cohen’s Kappa coefficient of this research was 0.801 which can be considered as very good agreement. Therefore, the overall agreement between LiDAR detected sinkholes and ground-truth sinkholes is very good. This proved that LiDAR data can be used to accurately detect and map the location of sinkholes.

Table 5. Cohen’s Kappa coefficient interpretation.

Cohen’s Kappa	Degree of Agreement
0.0 – 0.20	Poor
0.21 – 0.39	Fair
0.40 – 0.59	Moderate
0.60 – 0.79	Good
> 0.80	Very Good

5.4.2. Individual Level Validation

Individual level validation was focused on investigating if the LiDAR detected sinkholes and the ground-truth sinkholes have statistically similar morphometric measurements (i.e., surface area). In order to select the most appropriate method for comparing two groups of data, the population size of each sinkhole dataset (i.e., LiDAR detected sinkholes and ground-truth sinkholes) was examined. It revealed that the population size for both sinkhole datasets is 107, which is the TP value aforementioned. Most statistical research scientists accept that parametric statistical tests should be used if the sample size is equal to or greater than 30, unless the sample is not normally distributed. Therefore, the project team graphed the distribution of the two sinkhole datasets in Figures 30 to 32. They revealed that both datasets are not normally distributed.

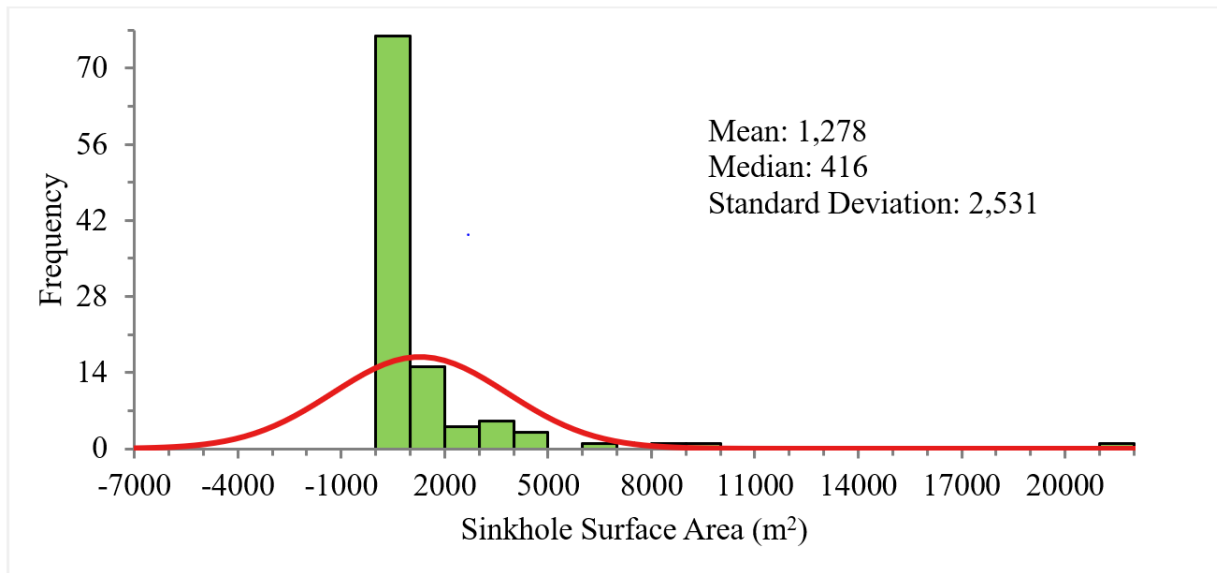


Figure 30. The distribution of the airborne LiDAR detected sinkhole surface area values.

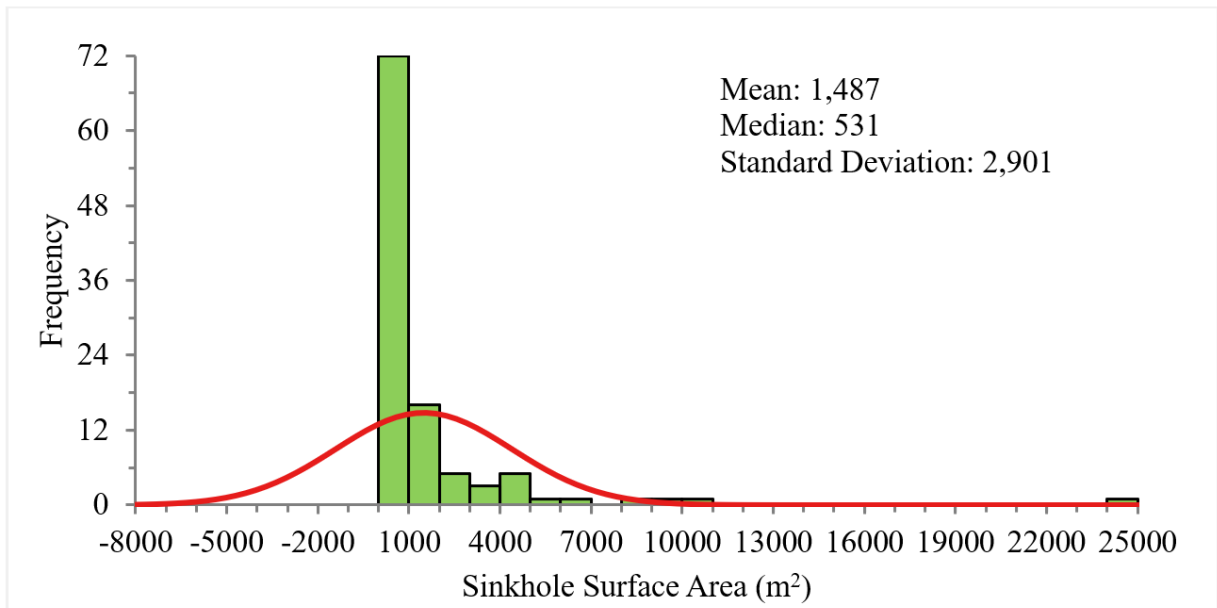


Figure 31. The distribution of the ground-truth sinkhole surface area values.

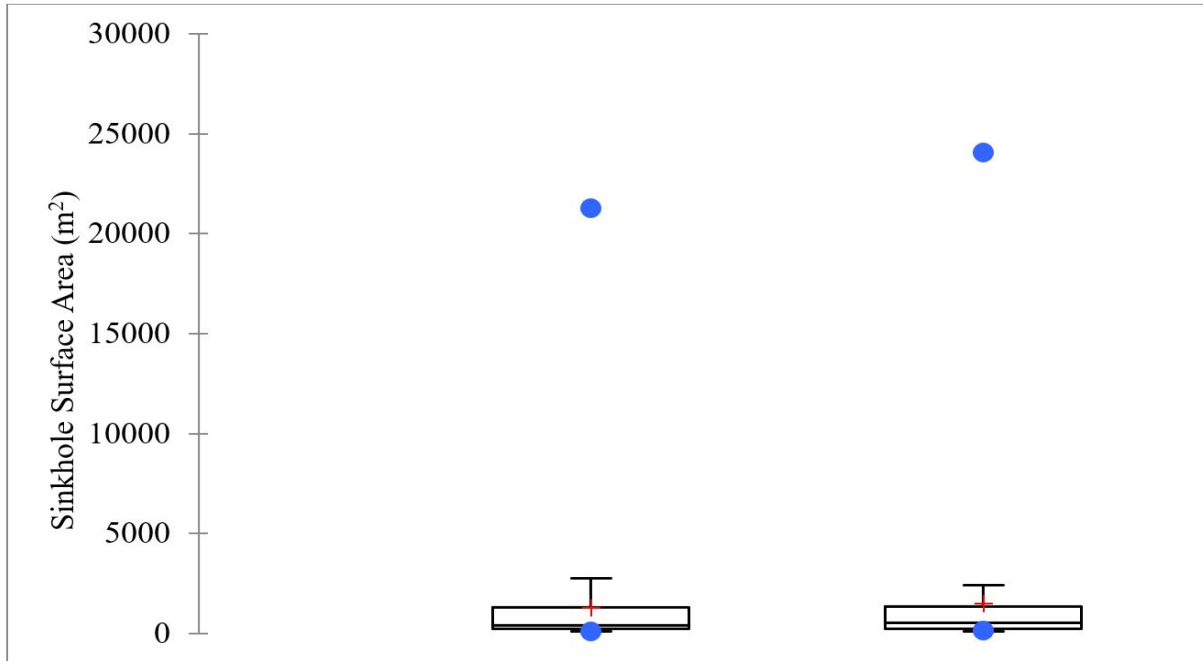


Figure 32. Box plot of the airborne LiDAR detected sinkhole and ground-truth sinkhole morphometric measurements.

Continuing with visual analysis, formal statistical tests were performed. Shapiro-Wilk test was used to test if the two sinkhole datasets are normally distributed. Table 5 shows the Shapiro-Wilk test results for both sinkhole datasets.

Table 6. Normality test results for LiDAR detected and ground-truth sinkhole layers.

Dataset	Null Hypothesis	P-value
LiDAR Detected Sinkholes	The distribution of the population is normal	< 0.0001
Ground-Truth Sinkholes	The distribution of the population is normal	< 0.0001

The normality test results revealed that both LiDAR detected sinkholes and ground-truth sinkholes are not normally distributed because the p values are less than 0.05. Therefore, nonparametric statistical tests were selected for comparing groups. Morphometric measurements were performed as a paired group and unpaired group. Paired group tests are more appropriate if two groups of measurements are dependent (i.e., repeated measurement for the same subject but at two different times). Unpaired group tests are more appropriate if two groups of measurements are independent (i.e., measurement for one sample in Group A has no bearing on the measurement for one sample in Group B). The relationship between LiDAR detected sinkhole measurements and ground-truth sinkhole measurements can be interpreted in both a dependent way and an independent way. In the dependent way, repeated measurements of a specific sinkhole were performed from LiDAR detected sinkholes and on the aerial photo digitized sinkholes. In the independent way, LiDAR detected sinkhole morphometric measurements have no bearing on the aerial photo based morphometric measurements since they are from two different data sources. Since the relationship can be interpreted in both ways, to err on the side of caution, this research used both paired group

and unpaired group statistical tests to examine if the morphometric measurements from LiDAR detected sinkholes and ground-truth sinkholes are statistically different.

In the paired group test, repeated measurements (one from LiDAR detected sinkholes while the other one from ground-truth sinkholes) of a specific sinkhole constitute a pair, and the purpose of this comparison is to examine if the median of differences between pairs of measurements is zero. Nonparametric Wilcoxon Signed Rank Test, which does not assume normality in the data, was used in this study as a robust alternative to parametric Student's T-Test. The null hypothesis is the median of differences between pairs of measurements is zero. The p-value for the test is less than 0.0001. Therefore, it is necessary to reject the null hypothesis and be in favor of the alternative hypothesis at the 0.05 significance level. That being said, the median of differences between pairs of measurements is not zero. Further exploration revealed that the LiDAR detected sinkhole boundaries sometimes are not completed or finished. In other words, sometimes there are holes within a sinkhole boundary. These holes also result in LiDAR detected sinkhole boundary being consistently smaller than that of the ground-truth sinkhole. Figure 33 explains the aforementioned issue. This is caused by no difference between the filled DTM's and original DTM's pixel values or simply by no pixel values (airborne LiDAR data collection missed this small area). On the other hand, the use of aerial photo digitization to create ground-truth sinkhole boundary could also result in incorrect delineation because sometimes it is very challenging to know where the boundary line falls. Therefore, it is not surprise to obtain a statistical result that is not significant. However, as it shows in Figure 33, the LiDAR detected sinkhole boundary and the ground-truth sinkhole boundary are very close. For all of the matched sinkholes, the project team did not find any one of them has the issue of boundary shift. Boundary shift issue is the LiDAR detected sinkhole boundary just barely touched the ground-truth sinkhole boundary.

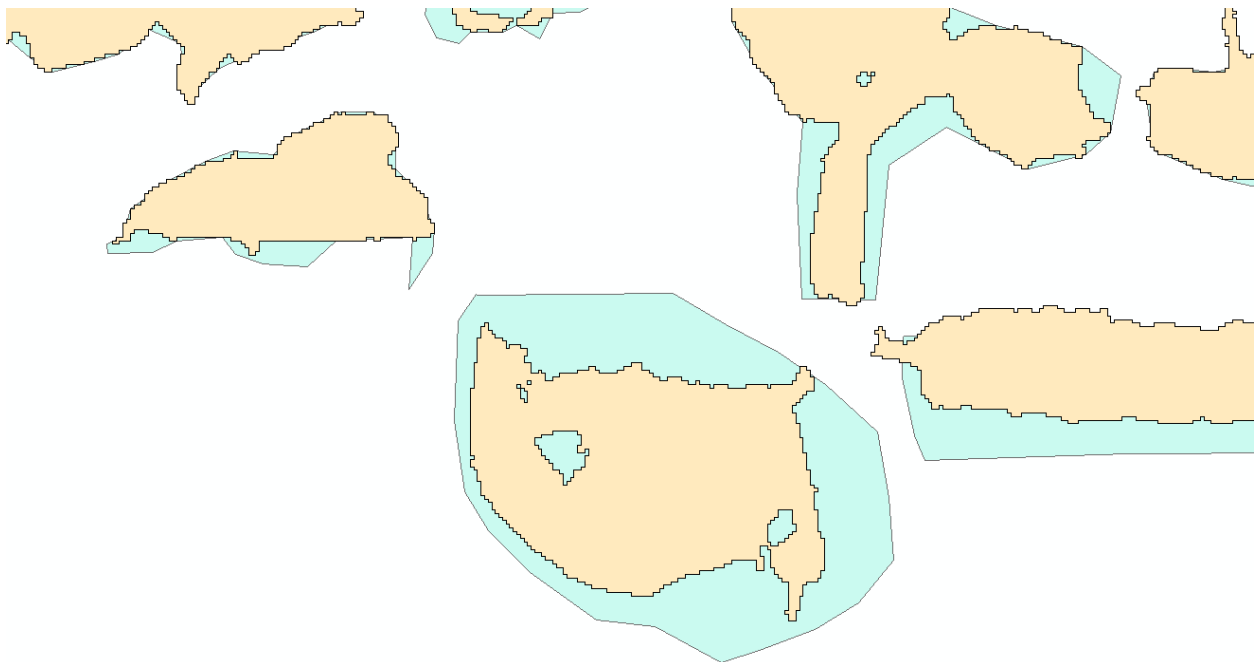


Figure 33. The boundary of LiDAR detected sinkholes (in beige color) and the ground-truth sinkholes (in jade color).

In the unpaired group test, two sets of morphometric measurements constitute two independent groups, and the purpose of this test is to examine whether two independent groups of measurements

exhibit the same distribution patterns (i.e., shape and spread) or have differences in medians. Nonparametric Mann-Whitney U test, also known as Wilcoxon Rank-Sum test, which also does not assume normality in the data, was used to detect differences in shape and spread as well as differences in medians. For this test, data from each population must be an independent random sample, and the population must have equal variances. For non-normally distributed data, the Levene's test is usually adopted to determine variance equability.

For the Levene's test, the null hypothesis is that the population variances are equal. The p-value for this test is 0.728, and therefore, the null hypothesis should be accepted. This indicated that the population variances for each group of comparison are equal at a 0.05 significance level. Therefore the Mann-Whitney U test is appropriate for comparing the differences in shape and spread as well as difference in medians. For the Mann-Whitney U test, the null hypothesis is that there is no difference in the distribution (shape and spread) of ground-truth morphometric measurements and LiDAR detected morphometric measurements. The p-value for this test is 0.282, and therefore, the null hypothesis should be retained, meaning there is no significant difference in the distribution pattern at a 0.05 significance level. This result revealed that at a group level, the LiDAR detected sinkhole morphometric measurements and the ground-truth measurements are statistically similar. This further proved that LiDAR can provide valid information that can be used to detect and delineate sinkhole boundary to generate a sinkhole database in a prompt fashion.

6. CONCLUSIONS

Sinkholes subsidence and collapse may cause significant structural damage to transportation infrastructure assets and are primarily a problem in karst areas where soluble bedrock is predominant. Subsequently, transportation infrastructure management agencies at all levels (e.g., federal, state, tribal, and local) dedicate a large amount of time and money to detect and map sinkholes as part of their infrastructure asset management programs. These collected sinkhole data are used by the aforementioned agencies to determine the extent and severity of sinkholes, and then to make decisions on the distribution of limited resources for sinkhole hazard mitigation to avoid potential public safety threats and problems.

Traditionally, sinkholes are primarily detected through area reconnaissance, which includes visual inspection of a site to locate existing sinkholes or device inspection of a site to locate potential sinkholes or previously filled sinkholes. This method is expensive, time-consuming, labor-intensive, and potentially dangerous to the inspectors. More importantly, because of its ground-based nature, this method is limited not only by the accessibility of the site but also the ability of the inspectors or devices to observe the entire site. For example, heavy-vegetation in a certain site may make it extremely difficult or even impossible to conduct any reconnaissance at all.

To overcome these challenges, this research focused on exploring the utility of airborne LiDAR in detecting and mapping sinkholes. This research used airborne LiDAR data in combination with not only object-based image analysis (OBIA) techniques but also auxiliary context information such as site and association to improve the accuracy of the current morphology-based sinkhole detection methods. Research results revealed that airborne LiDAR data derived DEMs can be used to identify all the possible sinkhole locations to develop preliminary sinkhole maps. The project team also successfully used OBIA and extract mask to conduct context-based feature extraction to delineate sinkhole boundaries and at the same time remove the false positively detected sinkholes. In addition, the project team used morphology-based feature extraction (i.e., sinkhole surface area) to further remove the false positively detected sinkholes.

The sinkhole map validation results have shown that the overall agreement between airborne LiDAR detected sinkholes and ground-truth sinkholes is very good. This proved that airborne LiDAR data can be used to accurately detect the location of sinkholes. With the help of the developed ArcGIS compatible toolset, this aforementioned sinkhole detection process can also be accomplished in a prompt fashion. At individual sinkhole level, the results revealed the airborne LiDAR detected sinkholes and ground-truth sinkholes do not have statistically similar morphometric measurements. This discrepancy were contributed by both the LiDAR detected sinkholes and ground-truth sinkholes, as discussed in the Analysis and Findings section. At individual group level, LiDAR detected sinkhole morphometric measurements and the ground-truth measurements are statistically similar. This further proved that airborne LiDAR can provide valid information that can be used to detect and delineate sinkhole boundary in a prompt fashion. The airborne LiDAR detected sinkholes can be potentially applied to evaluate overall sinkhole risks for rapid, high-level information checks.

It is revealed that the developed LiDAR sinkhole detection toolset is only limited by the spatial resolution and vertical accuracy of the input airborne LiDAR data. That being said, with higher spatial resolution and vertical accuracy input elevation data, the results are expected to be more effective. Therefore, future exploration should be focused on obtaining higher spatial resolution and vertical accuracy LiDAR data from various platforms, including, but not limited to, small

unmanned aircraft systems (sUAS), manned aircraft, and satellites. In addition, more accurate ground survey sinkhole data will be obtained to further investigate if the airborne LiDAR detected sinkholes and ground surveyed sinkholes have statistically similar morphometric measurements. In addition, this research did not focus on sinkhole depth because the project team did not have the depth information. However, the developed ArcGIS compatible toolset can be easily revised to extract the depth information from the LiDAR derived DEMs.

REFERENCES

1. Ford, D. C., and P. W. Williams. *Karst Geomorphology and Hydrology*. Unwin-Hyman, London, UK, 1989.
2. Miao, X., X. Qiu, S. S. Wu, J. Luo, D. R. Gouzie, and H. Xie. Developing Efficient Procedures for Automated Sinkhole Extraction from Lidar DEMs. *Photogrammetric Engineering & Remote Sensing*, 2013. 79(6): 545–554.
3. Zhu, J., T. P. Taylor, J. C. Currens, and M. M. Crawford. Improved Karst Sinkhole Mapping in Kentucky Using Lidar Techniques: A Pilot Study in Floyds Fork Watershed. *Journal of Cave and Karst Studies*, 2014. 76(3): 207–216.
4. McIntosh, C., Ed. *Cambridge Advanced Learner's Dictionary with CD-ROM*. Cambridge University Press, Cambridge, UK, 2013.
5. Tihansky, A. B. *Sinkholes, West-Central Florida - A Link Between Surface Water and Ground Water*. Publication 1182. U.S. Geological Survey, Reston, VA, 1999.
6. Pappas, S. What Are Sinkholes? *Live Science*. <https://www.livescience.com/44123-what-are-sinkholes.html>. Accessed Dec. 2, 2017.
7. Gouzie, D., and G. Pendergrass. Investigation and Remediation of the 2006 Nixa, Missouri, Collapse Sinkhole. *Environmental and Engineering Geoscience*, 2009. 15(1): 13–27.
8. Seale, L. D., L. J. Florea, H. L. Vacher, and R. Brinkmann. Using ALSM to Map Sinkholes in the Urbanized Covered Karst of Pinellas County, Florida—1, Methodological Considerations. *Environmental Geology*, 2008. 54(5): 995–1005.
9. Sevil, J., F. Gutiérrez, M. Zarroca, G. Desir, D. Carbonel, J. Guerrero, R. Linares, C. Roqué, and I. Fabregat. Sinkhole Investigation in an Urban Area by Trenching in Combination with GPR, ERT and High-Precision Leveling. Mantled Evaporite Karst of Zaragoza City, NE Spain. *Engineering Geology*, 2017. 231: 9–20.
10. Shapiro, S. D., E. Busenberg, and L. N. Plummer. *The Stability of Chlorofluorocarbons (CFCs) in Ground-Water Samples Archived in Borosilicate Ampoules*. Publication USGS OFR 2004-1392. U.S. Geological Survey, Reston, VA, 2005.
11. Weary, D. The Cost of Karst Subsidence and Sinkhole Collapse in the United States Compared with Other Natural Hazards. *Proceedings of the 14 Sinkhole Conference*, Rochester, Minnesota, 2015.
12. Filin, S., and A. Baruch. Detection of Sinkhole Hazards Using Airborne Laser Scanning Data. *Photogrammetric Engineering & Remote Sensing*, 2010. 76(5): 577–587.
13. Thomas, B., and M. J. S. Roth. Evaluation of Site Characterization Methods for Sinkholes in Pennsylvania and New Jersey. *Engineering Geology*, 1999. 52(1): 147–152.
14. Applegate, P. Detection of Sinkholes Developed on Shaly Ordovician Limestones, Hamilton County, Ohio, Using Digital Topographic Data: Dependence of Topographic Expression of Sinkholes on Scale, Contour Interval, and Slope. *Journal of Cave and Karst Studies*, 2003. 65(2): 126–129.

15. Wilson, K. V. *Modification of Karst Depressions by Urbanization in Pinellas County, Florida*. University of South Florida, Tampa, FL, 2004.
16. Ernst, C. L., J. Hadizadeh, and J. L. McCarty. A Remote Sensing Study of Regional Variation in Sinkhole Morphology-Florida Karst vs. Minnesota Karst. *Proceedings of the AGU 2010 Fall Meeting*, San Francisco, CA, 2010.
17. Hodgson, M. E., and P. Bresnahan. Accuracy of Airborne Lidar-Derived Elevation. *Photogrammetric Engineering & Remote Sensing*, 2004. 70(3): 331–339.
18. Arundel, S. T., L. A. Phillips, A. J. Lowe, J. Bobinmyer, K. S. Mantey, C. A. Dunn, E. W. Constance, and E. L. Usery. Preparing the National Map for the 3D Elevation Program – Products, Process and Research. *Cartography and Geographic Information Science*, 2015. 42(sup1): 40–53.
19. Wang, L., and H. Liu. An Efficient Method for Identifying and Filling Surface Depressions in Digital Elevation Models for Hydrologic Analysis and Modelling. *International Journal of Geographical Information Science*, 2007. 20(2): 193–213.
20. Liu, H., and L. Wang. Mapping Detention Basins and Deriving Their Spatial Attributes from Airborne LiDAR Data for Hydrological Applications. *Hydrological Processes*, 2008. 22(13): 2358–2369.
21. Vacher, H. L., L. Seale, L. Florea, and R. Brinkmann. Using ALSM to Map Sinkholes in the Urbanized Covered Karst of Pinellas County, Florida—2. Accuracy Statistics. *Environmental Geology*, 2008. 54(5): 1007–1015.
22. Filin, S., A. Baruch, Y. Avni, and S. Marco. Sinkhole Characterization in the Dead Sea Area Using Airborne Laser Scanning. *Natural Hazards*, 2011. 58(3): 1135–1154.
23. McHugh, M. L. Interrater Reliability: The Kappa Statistic. *Biochemia Medica*, 2012. 22(3): 276–282.

APPENDIX A: BEST PRACTICES FOR IMPLEMENTING A STATE-LEVEL SINKHOLE HAZARD MANAGEMENT SYSTEM (SHMS)

1. The SHMS should be developed based on a standard geographic information system (GIS).
2. The state transportation management agency should have specialized staff on a regular basis. This specialized staff should have a solid background in both sinkhole inspection and GIS.
3. The state transportation management agency needs to develop a complete database to understand the spatial distribution of the sinkholes.
4. When residents see a sinkhole appear, the state transportation management agency should immediately protect the surrounding areas so that nobody walks closer to the sinkhole.
5. The state transportation management agency should keep in touch with various state agencies to obtain the most current airborne LiDAR data or other elevation data that have been collected. And if at all possible, the state transportation management agency should also consider contributing funding to airborne LiDAR data collection.
6. When ground inspection performed for sinkholes, the field inspectors should take detailed photos and notes and this information should be updated in the sinkhole database as soon as the inspectors return back to the office.
7. The sinkhole geodatabase should use vector polygons instead of vector points. Two benefits will be provided by the polygon file. First, the morphometric measurements (e.g., length, width, area) will be automatically updated if the database administrator updated the database. Second, polygons can provide richer information regarding sinkholes than points. When the polygons are visualized on a map, the audience can easily understand the shape and distribution of the sinkholes.
8. The state transportation management agency should explore the utility of Volunteer Geographic Information (VGI) in obtaining real-time sinkhole status. VGI is a spatial data that has been contributed for free by citizen volunteers.
9. Some useful attributes should be considered in the SHMS, including, but not limited to, sinkhole ID, location, county, date identified, source of report, drainage status, shape, length, width, depth, property damage, feature repaired, witness name, witness phone, verified sinkhole, and comments.
10. The state transportation management agency should develop a web platform to allow the public to have free access to the sinkhole information. The web interface should be user-friendly and have the basic functions to allow the users to be able to examine the location of the sinkholes, the date that the sinkhole was identified, and the size and shape of the sinkholes.
11. The state transportation management agency should collaborate with local universities or colleges for workforce development and professional training.
12. The state transportation management agency should collaborate with National Cave and Karst Research Institute (NCKRI) to develop a complete working procedure to summarize the morphology characteristics of local sinkholes.

APPENDIX B: LIDAR BASED SINKHOLE DETECTION AND MAPPING TOOLSET GUIDEBOOK

This guidebook is developed to assist transportation management agencies with professional training and education and workforce development. This guidebook provides detailed instructions on how to use the developed ArcGIS compatible airborne LiDAR based sinkhole detection and mapping toolset.

Minimum Computer System Requirements:

- Memory: 8GB and more
- CPU: Intel Core i5
- GPU: 2GB and more
- Hard Drive: 512GB and more

Part A. Toolset Download

1. Go to the [download page](#); the download page looks like the following.

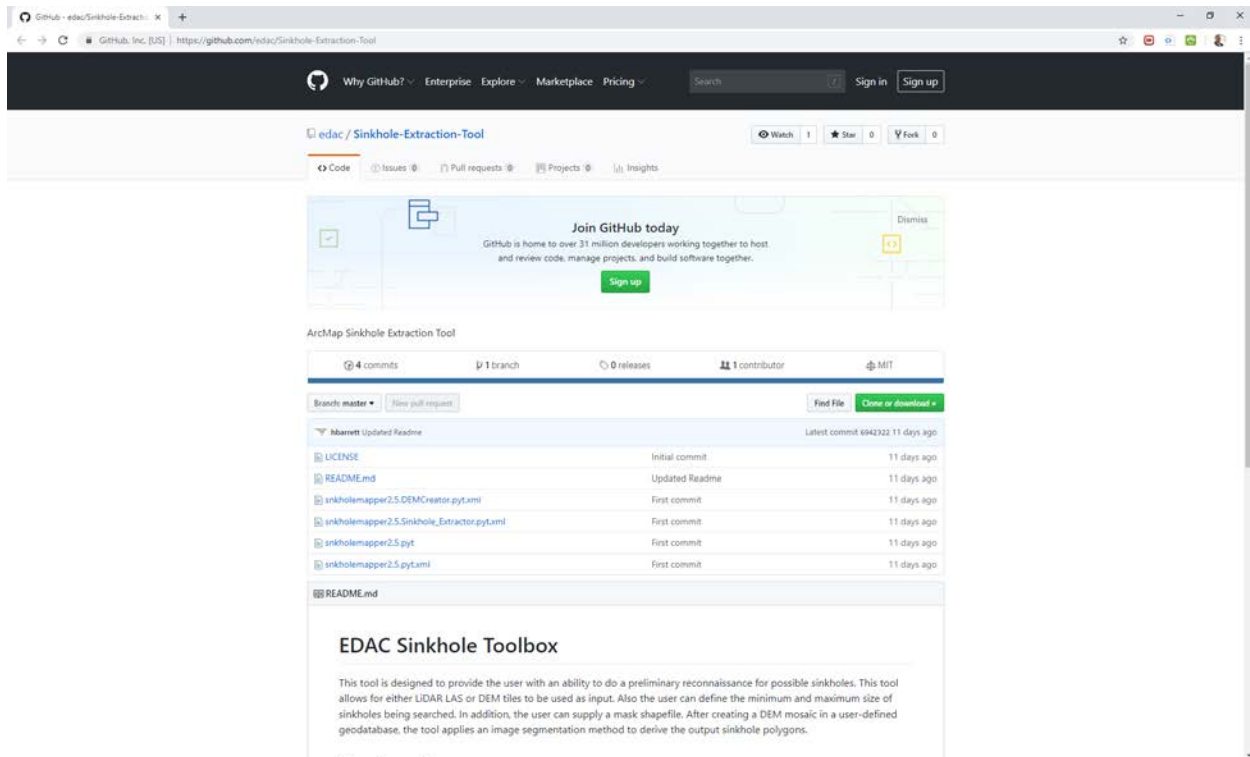


Figure B1. Downloading page interface for the sinkhole extraction tool.

2. Click on the button of “Clone or download”. This following dialog will pop up.

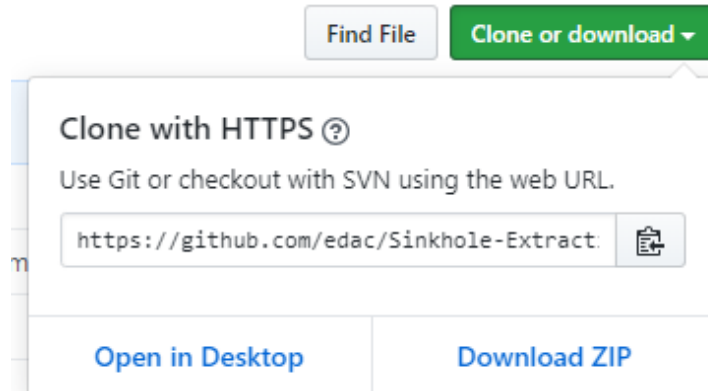


Figure B2. Clone and download button and its interface.

3. Click on Download ZIP. The download will start in a few seconds.
4. The users can find the toolset is compressed in a zip file once the download completed.

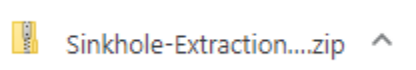


Figure B3. Downloaded sinkhole extraction tool.

Part B. Toolset Installation

1. Unzip the downloaded toolset.
2. Open ArcMap 10.3.1 or newer version.
3. The ArcMap user interface looks like the following.

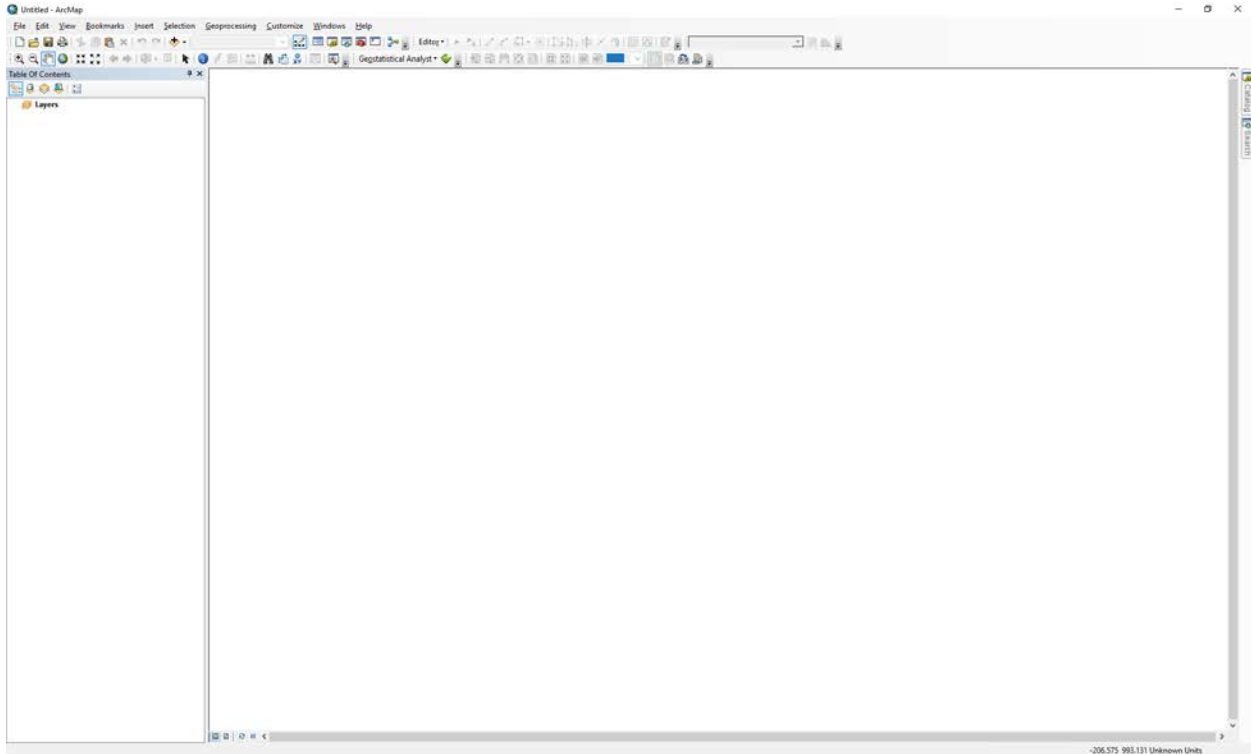


Figure B4. ArcMap software user interface.

3. In the Standard Toolbar, click on the ArcToolbox button below (the icon with the red box).

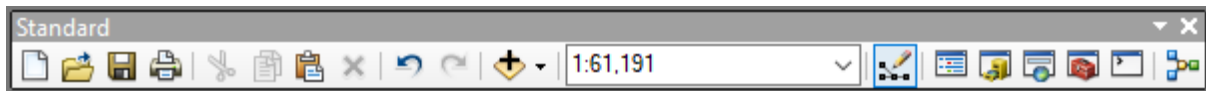


Figure B5. ArcMap software standard toolbar.

4. The ArcToolbox dialog will open. See the screen capture in the next page.

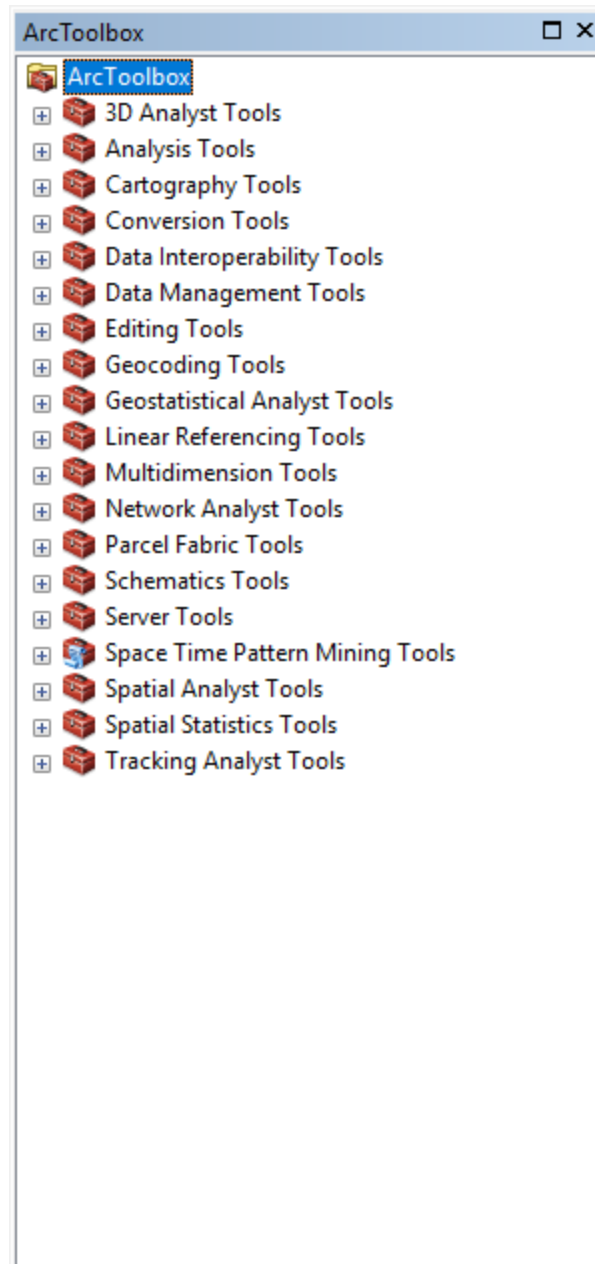


Figure B6. ArcMap ArcToolbox user interface.

5. Right-click on any blank space in the ArcToolbox, and then the Add Toolbox button will show up.

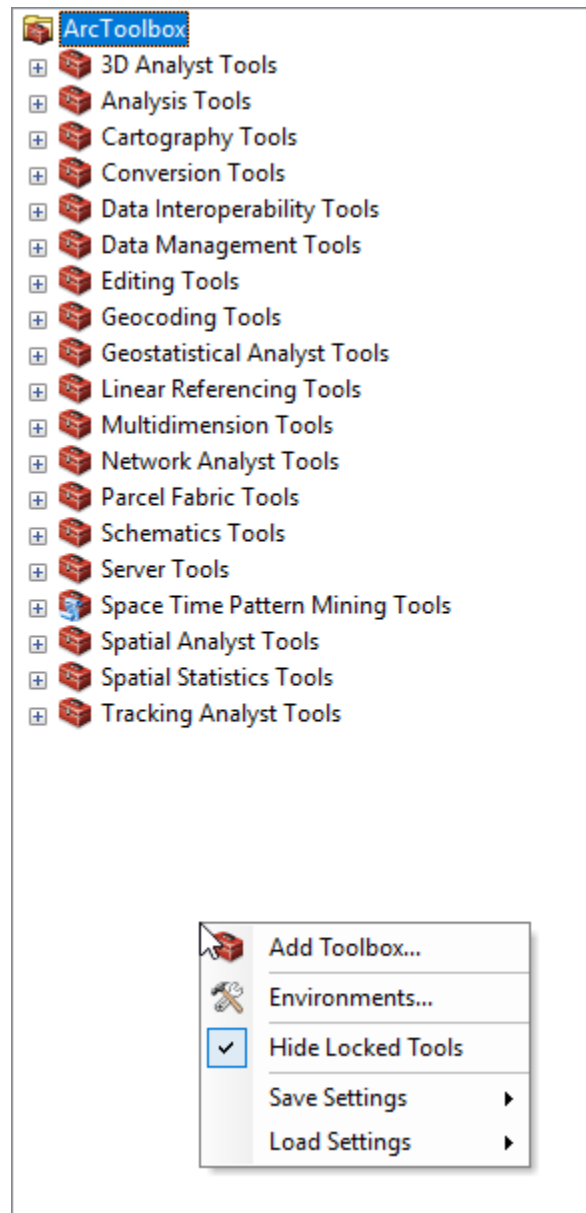


Figure B7. Add Toolbox button.

6. Click on the Add Toolbox button.

7. Browse to the folder where the sinkhole toolset was saved at.

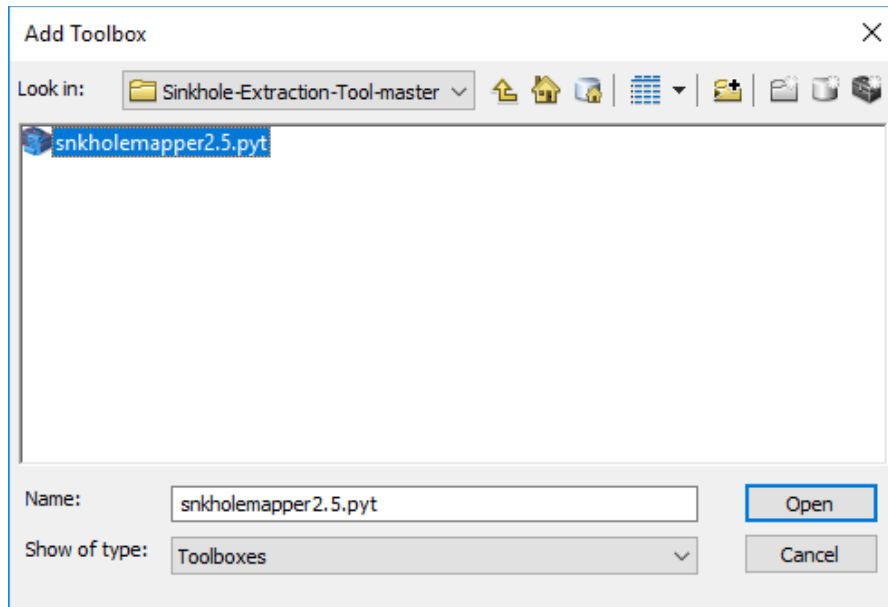


Figure B8. AddToolbox user interface.

8. Select the sinkhole mapper and then click on Open. The toolset will be added to the ArcToolbox.

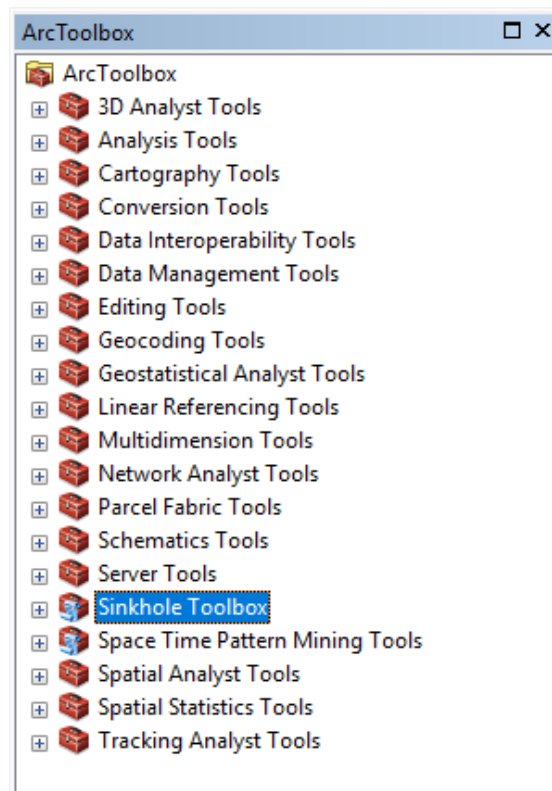


Figure B9. Added sinkhole toolbox.

9. Click on the plus sign next the Sinkhole Toolbox. The two tools associated with this toolset will show up.

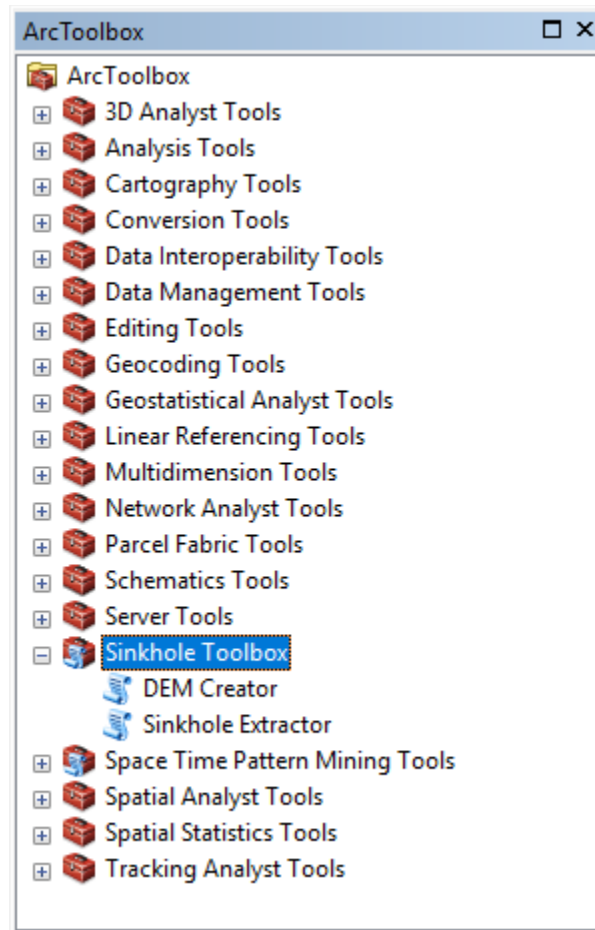


Figure B10. Expanded sinkhole toolbox.

10. Double-click on the DEM Creator tool. The user interface looks as the following screen capture.

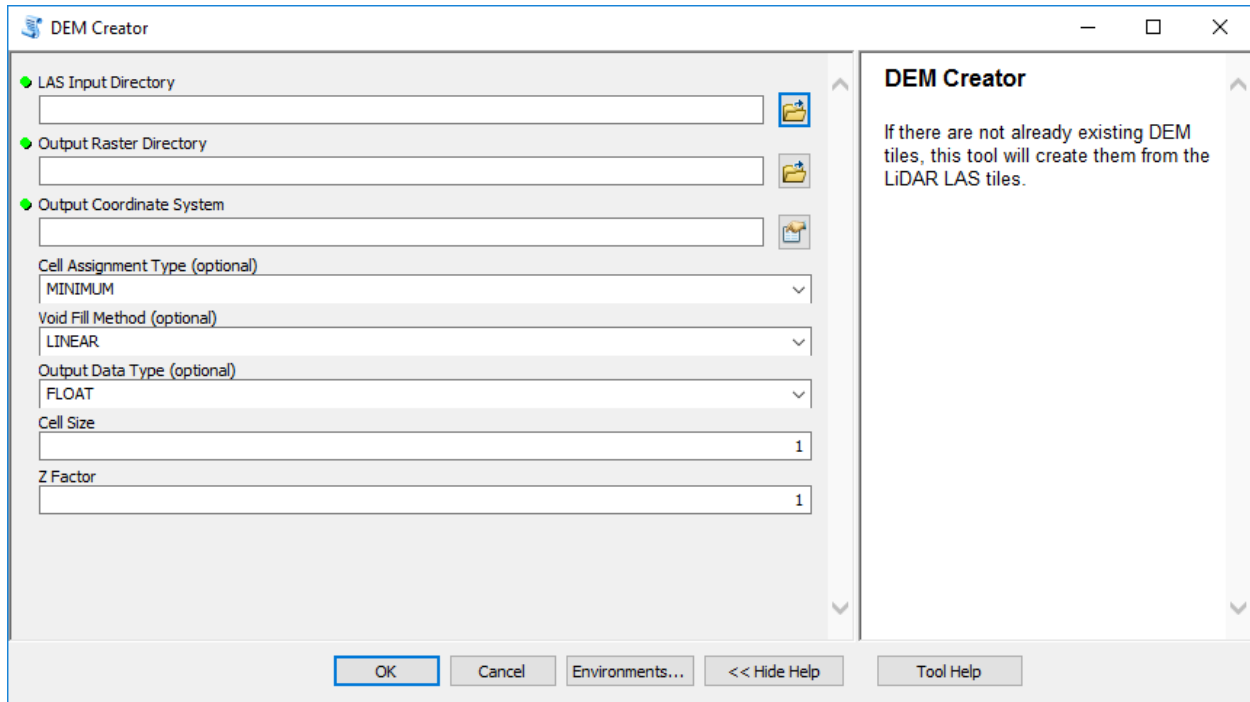


Figure B11. DEM creator tool user interface.

11. Double-click on the Sinkhole Extractor tool. The user interface looks as the following screen capture.

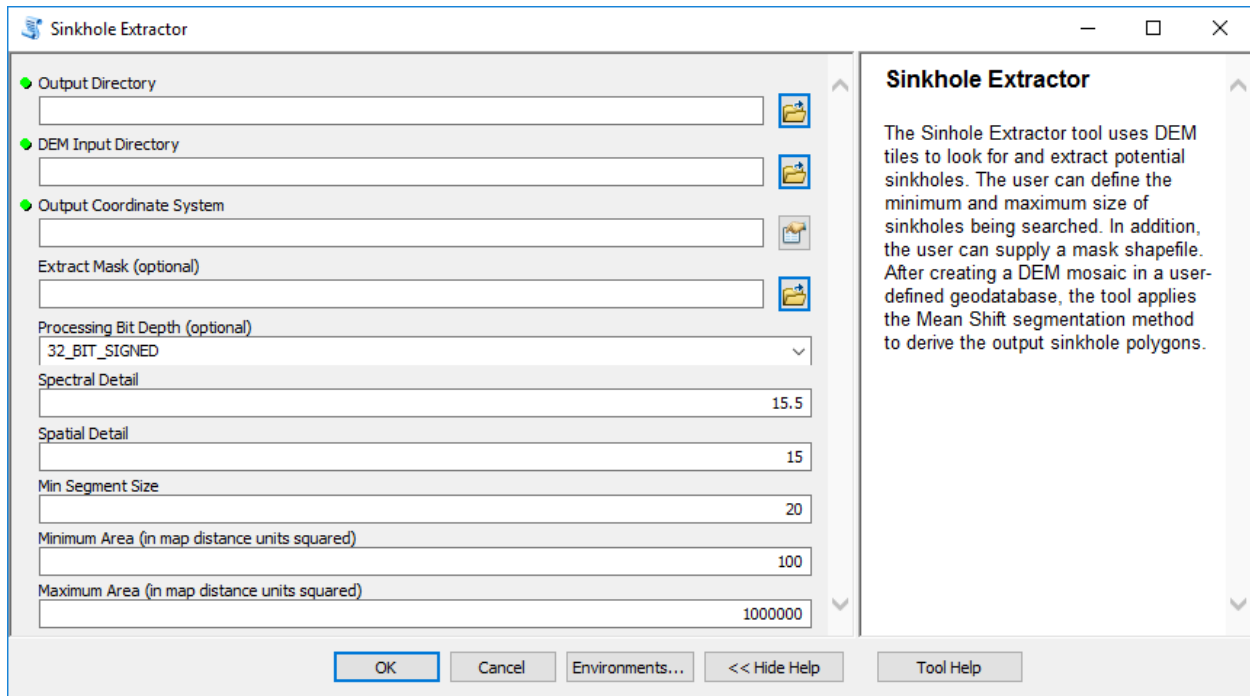


Figure B12. Sinkhole extractor tool user interface.

12. Users can click on the Hide Help to hide the tool instruction on the right column of the tools. Users can also click on the Show Help to display the tool instruction again.

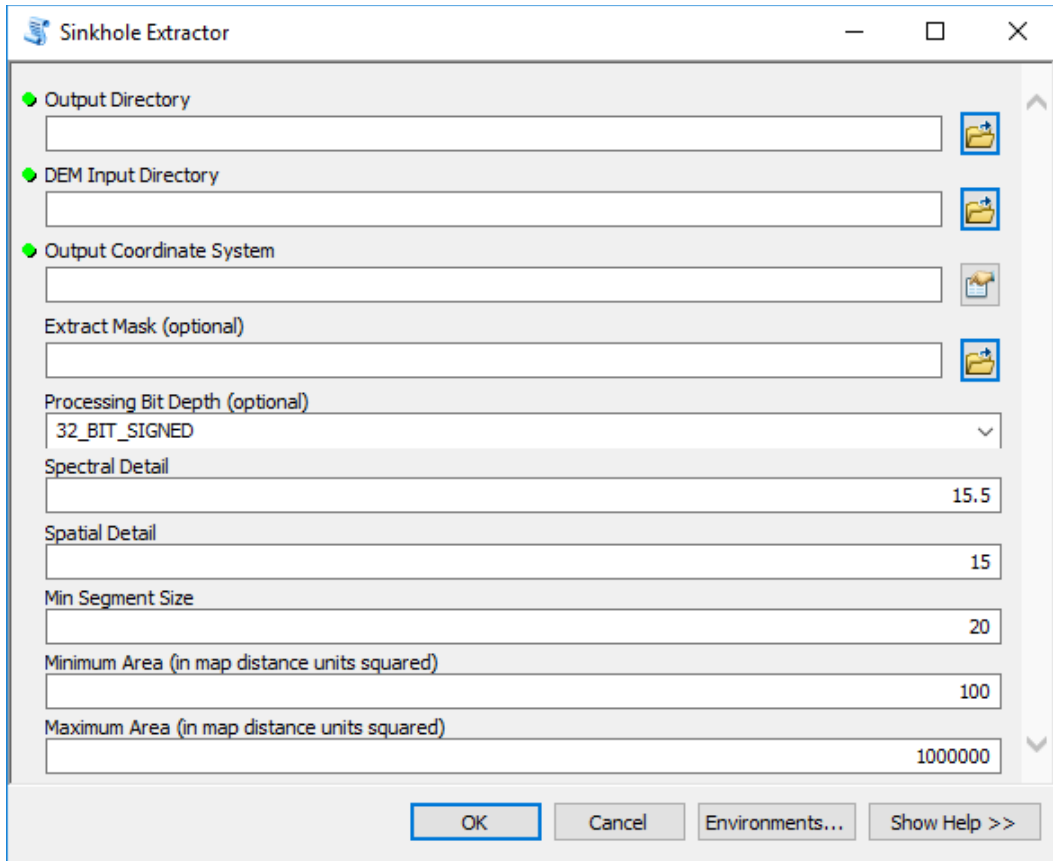


Figure B13. A way to close and open Help dialog.

Part C. Instructions on How to use the DEM Creator Tool

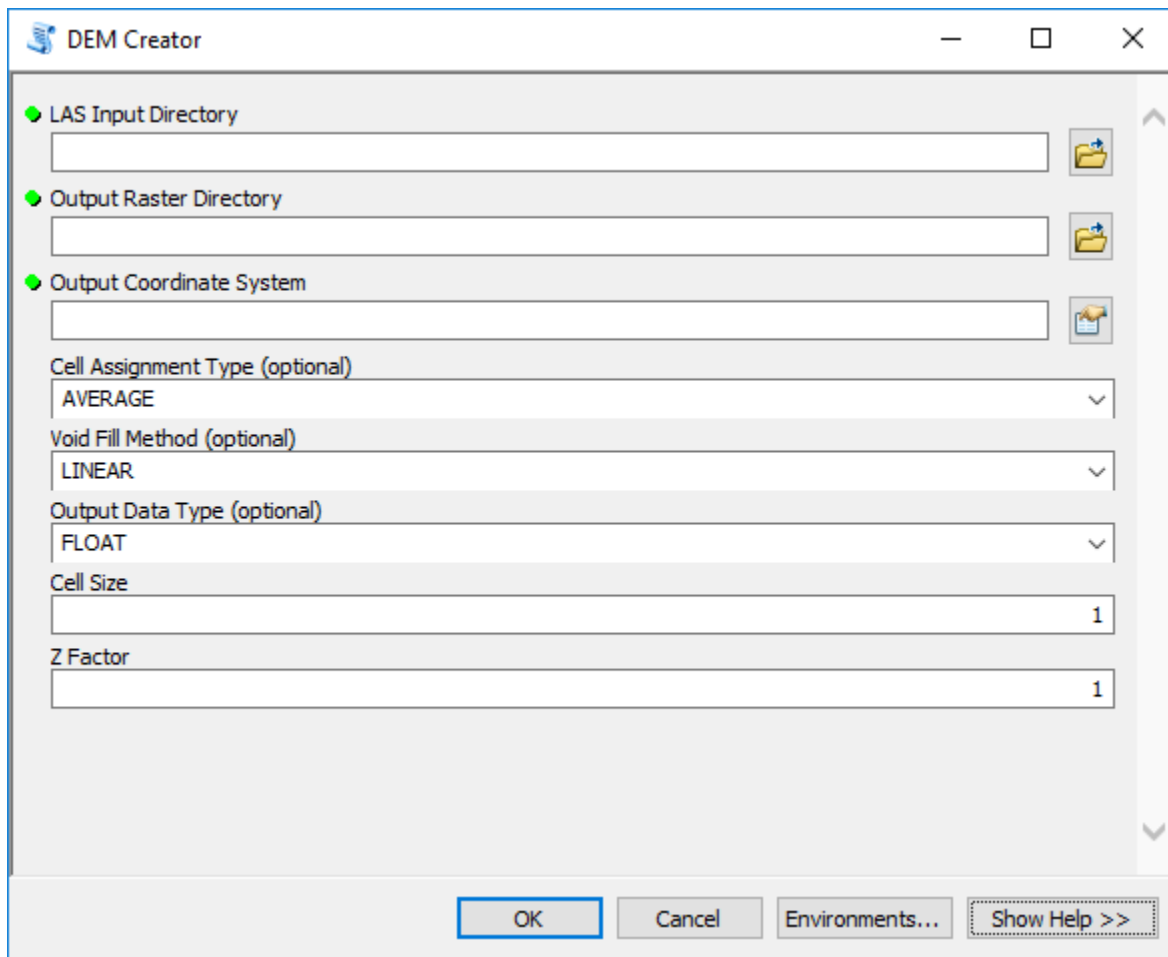


Figure B14. DEM creator tool with parameters filled in.

If the users only have the raw LiDAR LAS files, this tool can help the users create the needed digital elevation model (DEM) for the Sinkhole Extractor tool. There are eight parameters for this tool.

1. LAS Input Directory.

This is the directory where the input LiDAR LAS tiles are found.

2. Output Raster Directory

This is the directory in which the output DEM raster will be placed. The output DEM raster will be a single tile if the users input a single LAS tile. The output DEM raster will be also be single tile if the users input many LAS tiles, but the single DEM tile's spatial coverage will be equal to all input LAS tiles'.

3. Output Coordinate System

The coordinate system of the original DEM tiles. This needs to be in a projected coordinate system and not a spheroidal (geographic) coordinate system.

4. Cell Assignment Type (optional)

This is an optional parameter. Users are suggested to use the default value. Cell Assignment Type is used to determine each output cell's value using the LiDAR points that fall within its extent. Average will find the average value of all points within the cell and then assign the average value to the cell. Minimum will find the minimum value of all points within the cell and then assign the minimum value the cell. This is the default option. Maximum will find the maximum value of all points within the cell and then assign the maximum value the cell. IDW will find the inverse distance weighted interpolation value of all points within the cell and then assign the IDW value to the cell. Nearest will find the nearest neighbor LiDAR point to the cell centroid and then use that LiDAR point's value to determine the cell value.

5. Void Fill Method (optional)

This is an optional parameter. Users are suggested to use the default value. Void Fill Method is used to determine the values for cells that do not contain any LiDAR points. None – NoData is assigned to the cell. Simple – Averages the values from data cells immediately surrounding a NoData cell to eliminate small voids. Linear – Triangulates across void areas and users linear interpolation on the triangulated value to determine the cell value. This is the default. Natural Neighbor – Uses natural neighbor interpolation to determine where the void is and then assign that natural neighbor polygon's value to that void.

6. Output Data Type (optional)

This is the output raster type of the DEM. Users are suggested to use the default value. There are two options. FLOAT – The output will be floating point which will preserve the original elevation values. This is the default. INT –The output will be integer and truncate the original elevation value to nearest integer, but will create a smaller size file.

7. Cell Size

The size of the output cell in projection distance units.

8. Z Factor

The factor by which Z values will be multiplied. This is typically used to convert Z linear units to match XY linear units. The default is 1, which leaves elevation values unchanged.

Part D. Instructions on How to use the DEM Creator Tool

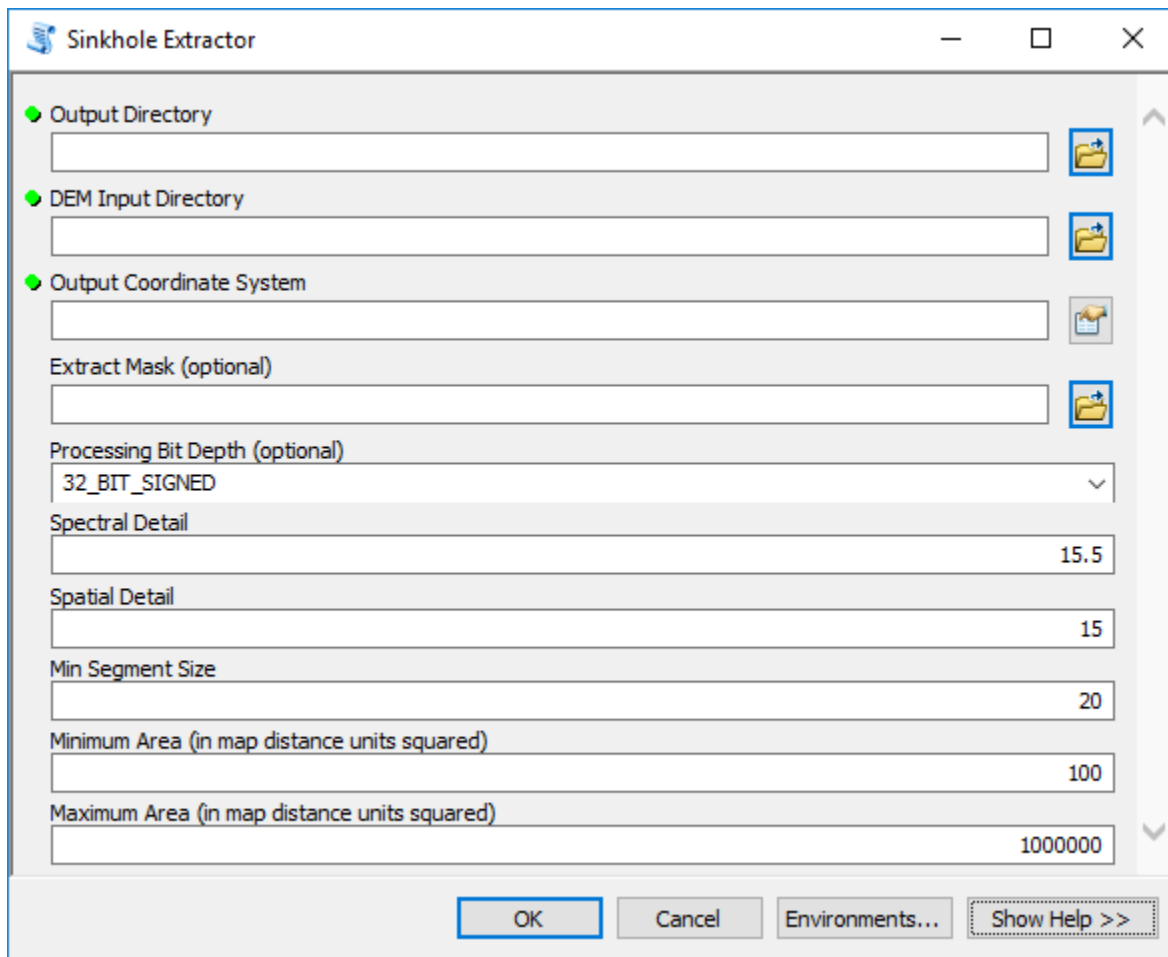


Figure B15. Sinkhole extractor tool with parameters filled in.

This tool uses DEMs tiles to look for and extract potential sinkholes. The user can define the minimum and maximum size of sinkholes being searched. In addition, the user can supply an extract mask shapefile. After creating a DEM mosaic in a user-defined geodatabase, the tool applies the Mean Shift segmentation method to derive the output sinkhole polygons. There are ten parameters for this tool.

1. Output Directory

This is the directory in which the output geodatabase will be placed. This geodatabase will have the DEM mosaic and the sinkhole polygon feature class.

2. DEM Input Directory

This is the directory where the DEM tiles are found.

3. Output Coordinate System

The coordinate system of the original DEM tiles. This needs to be in a projected coordinate system and not spheroidal (geographic) coordinate system.

4. Extract Mask (optional)

This is an optional parameter. It has to be shapefile and it will be used to constrain the analysis to those areas where sinkholes are expected. This should be in the same projection as the DEM data. The default is that the Extract Mask do not exist.

5. Processing Bit Depth (optional)

This is an optional parameter. This parameter is used to determine the output bit depth of the DEM mosaic. Selecting 32 Bit Signed option will truncate the elevation values to the nearest integer value which will output a smaller DEM mosaic file but will lose some elevation detail. This is the default. Selecting 32 Bit Float option will preserve the elevation values which will output a larger DEM mosaic but preserve the elevation detail. Selecting 64 Bit will preserve the elevation values which will output a larger DEM mosaic but preserve the elevation detail. This should be selected if the 32 Bit Float option cannot provide enough value range.

6. Spectral Detail

During the image segmentation process this sets the level importance given to the spectral differences of features in the imagery. A default value is set to 15.5 but valid values range from 1.0 to 20.0. A higher value is appropriate when the features are spectrally similar. Smaller values create smoother outputs. If the output does not identify enough features, then raise the value. If it identifies too many features then lower the value.

7. Spatial Detail

During the image segmentation process this sets the level importance given to the spatial differences of features in the imagery. A default value is set to 15.5 but valid values range from 1.0 to 20.0. A higher value is appropriate when the features are spatially similar. Smaller values create smoother outputs. If the output does not identify enough features, then raise the value. If it identifies too many features then lower the value.

8. Min Segment Size

The minimum size of the raster object that will be output from the image segmentation process. Area is measured in the square distance units.

9. Minimum Area

The minimum size of the polygons that will be found in the output file Area is measured in the square distance units.

10. Maximum Area

The maximum size of the polygons that will be found in the output file Area is measured in the square distance units.

APPENDIX C. LIDAR BASED SINKHOLE DETECTION AND MAPPING TOOLSET CODES

```
import arcpy

import os

import datetime

from arcpy import env

from time import sleep

from arcpy.sa import *

import glob

import arceditor

arcpy.env.overwriteOutput = True

timestamp = datetime.datetime.now()
```

```

class Toolbox(object):

    def __init__(self):

        self.label = "Sinkhole Toolbox"

        self.alias = "EDAC Sinkhole Toolbox"

        # List of tool classes associated with this toolbox

        self.tools = [Sinkhole_Extractor,DEMCreator]

class DEMCreator(object):

    def __init__(self):

        self.label = "DEM Creator"

        self.description = "This tool creates DEM tiles from the LAS tiles which can be used for the
Sinkhole Extractor tool if bare earth DEM tiles are not available."

        self.canRunInBackground = False

    def getParameterInfo(self):

        lasdir = arcpy.Parameter(displayName="LAS Input Directory", name="LAS Input Directory",
datatype="DEFolder", parameterType="Required", direction="Input")

        outputraster = arcpy.Parameter(displayName="Output Raster Directory", name="Output
Raster Directory", datatype="DEFile", parameterType="Required", direction="Output")

        crs = arcpy.Parameter(displayName="Output Coordinate System", name="Output Coordinate
System", datatype="GPCoordinateSystem", parameterType="Required", direction="Input")

        # return_values = arcpy.Parameter(displayName="Return Values",name="Return
Values",datatype="GPString",parameterType="Optional",direction="Input")

        # return_values.value = "Last Return"

        # return_values.filter.type = "ValueList"

        # return_values.filter.list = ["Last Return","First of Many","Last of Many","Single
Return","1","2","3","4","5"]

```

```

    binningcell = arcpy.Parameter(displayName="Cell Assignment Type",name="Cell
Assignment Type",datatype="GPString",parameterType="Optional",direction="Input")

    binningcell.value = "MINIMUM"

    binningcell.filter.type = "ValueList"

    binningcell.filter.list = ["AVERAGE","MINIMUM","MAXIMUM","IDW","NEAREST"]

    binningvoid = arcpy.Parameter(displayName="Void Fill Method",name="Void Fill
Method",datatype="GPString",parameterType="Optional",direction="Input")

    binningvoid.value = "LINEAR"

    binningvoid.filter.type = "ValueList"

    binningvoid.filter.list = ["NONE","SIMPLE","LINEAR","NATURAL_NEIGHBOR"]

    outputdatatype = arcpy.Parameter(displayName="Output Data Type",name="Output Data
Type",datatype="GPString",parameterType="Optional",direction="Input")

    outputdatatype.value = "FLOAT"

    outputdatatype.filter.type = "ValueList"

    outputdatatype.filter.list = ["FLOAT","INT"]

    cell_size = arcpy.Parameter(displayName="Cell Size", name="Cell Size",
datatype="GPDDouble",parameterType="Required", direction="Input" )

    cell_size.value = 1

    z_factor = arcpy.Parameter(displayName="Z Factor", name="Z Factor",
datatype="GPDDouble",parameterType="Required", direction="Input" )

    z_factor.value = 1

    parameters =
[lasdir,outputtraster,crs,binningcell,binningvoid,outputdatatype,cell_size,z_factor]

    return parameters

def isLicensed(self): # optional

    return True

def updateMessages(self, parameters):

```

```
return
```

```
def execute(self, parameters, messages):
```

```
    lasdir = parameters[0].valueAsText
```

```
    outputraster = parameters[1].valueAsText
```

```
    crs = parameters[2].valueAsText
```

```
    return_values = "Last Return" #parameters[3].valueAsText
```

```
    binningcell=parameters[3].valueAsText
```

```
    binningvoid=parameters[4].valueAsText
```

```
    outputdatatype=parameters[5].valueAsText
```

```
    cell_size=parameters[6].valueAsText
```

```
    z_factor=parameters[7].valueAsText
```

```
    class_code="2"
```

```
    #define output paths
```

```
    temp=os.environ.get("TEMP")
```

```
    lasd=os.path.join(temp,"tempLASD.lasd")
```

```
    lasd2=os.path.join(temp,"tempLASD2.lasd")
```

```
# Execute CreateLasDataset
```

```
    arcpy.AddMessage("Execute Create Las Dataset")
```

```
    arcpy.management.CreateLasDataset(lasdir,lasd,spatial_reference=crs)
```

```

# Execute Make LAS Dataset Layer

arcpy.AddMessage("Make LAS Dataset Layer")

if class_code is not None and return_values is not None:

arcpy.MakeLasDatasetLayer_management(lasd,lasd2,class_code=class_code,return_values=return_values)

elif class_code is not None:

    arcpy.MakeLasDatasetLayer_management(lasd,lasd2,class_code=class_code)

elif return_values is not None:

    arcpy.MakeLasDatasetLayer_management(lasd,lasd2,return_values=return_values)

else:

    arcpy.MakeLasDatasetLayer_management(lasd,lasd2)

# Execute Las Dataset To Raster

# LasDatasetToRaster_conversion (in_las_dataset, out_raster, {value_field},
{interpolation_type}, {data_type}, {sampling_type}, {sampling_value}, {z_factor})

arcpy.AddMessage("Execute Las Dataset To Raster")

arcpy.LasDatasetToRaster_conversion (lasd2, outputraster,"ELEVATION", "BINNING
"+binningcell+" "+binningvoid , outputdatatype, "CELLSIZE", cell_size, z_factor)

```



```

class Sinkhole_Extractor(object):

    def __init__(self):

        self.label = "Sinkhole Extractor"

        self.description = "This tool extracts sinkholes"

        self.canRunInBackground = False

    def getParameterInfo(self):

        # Input parameters

        outputdir = arcpy.Parameter(displayName="Output Directory", name="Output Directory",
        datatype="DEFolder", parameterType="Required", direction="Input")

        demdir = arcpy.Parameter(displayName="DEM Input Directory", name="DEM Input
        Directory", datatype="DEFolder", parameterType="Required", direction="Input")

        crs = arcpy.Parameter(displayName="Output Coordinate System", name="Output Coordinate
        System", datatype="GPCoordinateSystem", parameterType="Required", direction="Input")

        mask = arcpy.Parameter(displayName="Extract Mask", name="Extract Mask",
        datatype="DEFeatureClass", parameterType="Optional", direction="Input")

        proc_bit_depth = arcpy.Parameter(displayName="Processing Bit Depth", name="Processing
        Bit Depth", datatype="GPString", parameterType="Optional", direction="Input")

        proc_bit_depth.value = "32_BIT_SIGNED"

        proc_bit_depth.filter.type = "ValueList"

        proc_bit_depth.filter.list = ["32_BIT_SIGNED", "32_BIT_FLOAT", "64_BIT"]

        spectral_detail = arcpy.Parameter(displayName="Spectral Detail", name="Spectral Detail",
        datatype="GPDDouble", parameterType="Required", direction="Input")

```

```

spectral_detail.value = 15.5

spatial_detail = arcpy.Parameter(displayName="Spatial Detail", name="Spatial Detail",
datatype="GPLong", parameterType="Required", direction="Input")

spatial_detail.value = 15

min_segment_size = arcpy.Parameter(displayName="Min Segment Size", name="Min
Segment Size", datatype="GPLong", parameterType="Required", direction="Input")

min_segment_size.value = 20

shapeareaMin = arcpy.Parameter(displayName="Minimum Area (in map distance units
squared)", name="Minimum Area", datatype="GPLong", parameterType="Required",
direction="Input")

shapeareaMin.value=100

shapeareaMax = arcpy.Parameter(displayName="Maximum Area (in map distance units
squared)", name="Maximum Area", datatype="GPLong", parameterType="Required",
direction="Input")

shapeareaMax.value=1000000

parameters = [outputdir, demdir, crs, mask, proc_bit_depth,spectral_detail, spatial_detail,
min_segment_size, shapeareaMin,shapeareaMax]

return parameters

def isLicensed(self): # optional

return True

def updateMessages(self, parameters):

return

def execute(self, parameters, messages):

```

```
arcpy.SetProgressor("default", "Working...", 0, 2, 1)
outputdir = parameters[0].valueAsText
dem_tiles = parameters[1].valueAsText
crs = parameters[2].valueAsText
mask = parameters[3].valueAsText #mask
proc_bit_depth = parameters[4].valueAsText
spectral_detail = parameters[5].valueAsText
spatial_detail = parameters[6].valueAsText
min_segment_size = parameters[7].valueAsText
shapeareaMin = parameters[8].valueAsText
shapeareaMax = parameters[9].valueAsText
```

Local variables:

```
sinkhole_gdb = os.path.join(outputdir, "Sinkhole.gdb")
Mosaic_Dataset = os.path.join(outputdir, "Sinkhole.gdb", "dem_mosaic")
Mosaic_Dataset_Final = os.path.join(outputdir, "Sinkhole.gdb", "dem_mosaic_final")
sinkholeDissolve = os.path.join(outputdir, "Sinkhole.gdb", "sinkholeDissolve")
Sinkhole_polygons = os.path.join(outputdir, "Sinkhole.gdb", "Sinkhole_polygons")
dem_fill_img = os.path.join(outputdir, "dem_fill.img")
dem_diff_img = os.path.join(outputdir, "dem_diff.img")
dem_diff2_img = os.path.join(outputdir, "dem_diff2.img")
dem_diff2mask_img = os.path.join(outputdir, "dem_diff2mask.img")
dem_diff2_is_img = os.path.join(outputdir, "dem_diff2_is.img")
sinkhole_polys_shp = os.path.join(outputdir, "sinkhole_polys.shp")
SegmentMeanShiftInput = None
```

```
# Process: Create output file geodatabase
arcpy.AddMessage("Creating file GDB")
arcpy.CreateFileGDB_management(outputdir, "Sinkhole.gdb")
productname=arcpy.GetInstallInfo()['ProductName']
arcpy.AddMessage(productname)
```

```
# Process: Create Mosaic Dataset
```

```
    arcpy.AddMessage("Creating Mosaic Dataset")
```

```
    arcpy.CreateMosaicDataset_management(sinkhole_gdb, "dem_mosaic", crs, "",  
proc_bit_depth, "NONE", "")
```

```
# Process: Add Rasters To Mosaic Dataset
```

```
    arcpy.AddMessage("Adding files from "+str(dem_tiles)+" to mosaic dataset.")
```

```
    #arcpy.AddRastersToMosaicDataset_management(Mosaic_Dataset, "Raster Dataset",  
dem_tiles, "UPDATE_CELL_SIZES", "UPDATE_BOUNDARY", "NO_OVERVIEWS", "",  
"0", "1500", "", "", "SUBFOLDERS", "ALLOW_DUPLICATES", "BUILD_PYRAMIDS",  
"CALCULATE_STATISTICS", "NO_THUMBNAILS", "",  
"NO_FORCE_SPATIAL_REFERENCE", "ESTIMATE_STATISTICS", "")
```

```
    arcpy.AddRastersToMosaicDataset_management(Mosaic_Dataset, "Raster Dataset",  
dem_tiles)
```



```
# Process: Set Null
```

```
arcpy.AddMessage("Setting any value less than 0 to Null for Mosaic Dataset.")
```

```
OutSetNull=SetNull(Mosaic_Dataset, Mosaic_Dataset, "VALUE <= 0")
```

```
OutSetNull.save(Mosaic_Dataset_Final)
```

```
# Process: Fill  
arcpy.AddMessage("Running fill on mosaic dataset.")  
outFill=Fill(Mosaic_Dataset_Final, "")  
outFill.save(dem_fill_img)
```

```
# Process: Raster Calculator
```

```
arcpy.AddMessage("Subtracting mosaic from mosaic fill.")
```

```
arcpy.Minus_3d(dem_fill_img, Mosaic_Dataset_Final, dem_diff_img)
```

```
# Process: Set Null
```

```
arcpy.AddMessage("Setting any value less than 0 to Null for resulting raster values.")
```

```
OutSetNull=SetNull(dem_diff_img, dem_diff_img, "VALUE <= 0")
```

```
OutSetNull.save(dem_diff2_img)
```

```
# Process: Extract by Mask if one is given
```

```
if mask is not None:
```

```
    arcpy.AddMessage("Mask is provided, so extracting data by mask.")
```

```
    ExtractBy=ExtractByMask(dem_diff2_img, mask)
```

```
    ExtractBy.save(dem_diff2mask_img)
```

```
    SegmentMeanShiftInput=dem_diff2mask_img
```

```
else:
```

```
    arcpy.AddMessage("No mask provided by user.")
```

```
    SegmentMeanShiftInput=dem_diff2_img
```

```
# Process: Segment Mean Shift
```

```
    arcpy.AddMessage("Grouping adjacent pixels together that have similar spectral  
characteristics.")
```

```
    SegMeShOut=SegmentMeanShift(SegmentMeanShiftInput, spectral_detail, spatial_detail,  
min_segment_size)
```

```
    SegMeShOut.save(dem_diff2_is_img)
```

```
# Process: Raster to Polygon
```

```
    arcpy.AddMessage("Saving Raster to polygon.")
```

```
    arcpy.RasterToPolygon_conversion(dem_diff2_is_img, sinkhole_polys_shp,  
    "NO_SIMPLIFY", "Value")
```



```
# Process: Write Polygon to sinkhole geodatabase
arcpy.AddMessage("Writing Polygon to GDB.")
arcpy.FeatureClassToGeodatabase_conversion(sinkhole_polys_shp, sinkhole_gdb)
```

```
# Process: Dissolve the polygons
```

```
    arcpy.AddMessage("Dissolving polygons.")
```

```
    arcpy.Dissolve_management(sinkhole_polys_shp, sinkholeDissolve, "", "",  
"SINGLE_PART", "")
```

Process: Select tool to select and output all polygons GTE 100 sq meters(or value provided by user).

```
arcpy.AddMessage("Selecting area greater than or equal to "+str(shapeareaMin)+" and less than or equal to "+ str(shapeareaMax) +" sq meters.")
```

```
where="Shape_Area >= "+ shapeareaMin + " AND " + "Shape_Area <= "+ shapeareaMax
```

```
arcpy.Select_analysis(sinkholeDissolve, Sinkhole_polygons, where)
```

```
# Process: Clean up extraneous files

arcpy.AddMessage("Cleaning up extraneous files")

arcpy.Delete_management(dem_fill_img)

arcpy.Delete_management(dem_diff_img)

arcpy.Delete_management(dem_diff2_img)

arcpy.Delete_management(dem_diff2mask_img)

arcpy.Delete_management(dem_diff2_is_img)

arcpy.Delete_management(sinkhole_polys_shp)

arcpy.Delete_management(os.path.join(outputdir, "Sinkhole.gdb", "sinkhole_polys"))

arcpy.Delete_management(sinkholeDissolve)
```

```
#openthis=os.path.join(outdir, "Sinkhole.gdb", "Sinkhole_polygons")

if productname=="ArcGISPro":

    arcpy.AddMessage("Will not display output in ArcGIS Pro.")

elif productname=="Desktop":

    mxd = arcpy.mapping.MapDocument("CURRENT")
    arcpy.mapping.ListDataFrames(mxd)[0].name = "Sinkhole Output"
    addLayer = arcpy.mapping.Layer(Sinkhole_polygons)
    df = arcpy.mapping.ListDataFrames(mxd)[0]
    arcpy.mapping.AddLayer(df, addLayer, "BOTTOM")
```