SINKHOLE FORMATION DYNAMICS AND GEOSTATISTICAL-BASED PREDICTION ANALYSIS IN A MANTLED KARST TERRAIN

by

MATTHEW DAVID CAHALAN

(Under the Direction of Adam Milewski)

ABSTRACT

This thesis presents an integrated, effective approach to understanding sinkhole formation dynamics and provides a decision-making tool to mitigate risks associated with sinkhole formation. Comprehensive ordinary least squares (OLS) and geographically weighted regression (GWR) geostatistical models are developed for quantifying sinkhole formation mechanisms in the mantled karst terrain of Dougherty County, Georgia (area: 183 km^2). Sinkhole density was determined using a GIS-based sinkhole mapping procedure with ten-meter resolution digital elevation models (DEMs) from 1999 and 2011 and a one-meter 2011 LiDAR DEM. Geostatistical models were performed on two sinkhole density datasets: 1) a spatiotemporal dataset representing newly formed and enlarged sinkholes between 1999 and 2011, and 2) LiDAR-derived sinkholes. Geostatistical results show that geologic, hydrologic, and hydrogeologic variables are most influential in sinkhole formation. Geostatistical prediction results were used with deterministic interpolators to assess the ability of statistically-based methods to predict sinkhole density.

INDEX WORDS: Sinkhole mapping, Sinkhole formation mechanism, LiDAR, Spatial statistics, GIS

SINKHOLE FORMATION DYNAMICS AND GEOSTATISTICAL-BASED PREDICTION ANALYSIS IN A MANTLED KARST TERRAIN

by

MATTHEW DAVID CAHALAN

B.A., University of Colorado Colorado Springs, 2013

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2015

SINKHOLE FORMATION DYNAMICS AND GEOSTATISTICAL-BASED PREDICTION ANALYSIS IN A MANTLED KARST TERRAIN

by

MATTHEW DAVID CAHALAN

Major Professor: Adam Milewski

Committee: Todd Rasmussen
Robert Hawman

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

ACKNOWLEDGEMENTS

I would like to extend my utmost appreciation to Dr. Adam Milewski for his tireless efforts supporting my education and continuously granting me opportunities to improve my scientific skill set and gain experience in a variety of areas.

I would like to thank Dr. Todd Rasmussen for improving my conceptual model of sinkhole formation and helping me to identify the limitations of the spatial statistics applied in this work. I am grateful to Dr. Rob Hawman for supporting this work through his study area field experience and sinkhole formation model development.

This project was made possible by the extensive data available from the U.S. Geological Survey and Georgia GIS Clearinghouse. I would like to thank Randy Weathersby from the City of Albany and Dougherty County Engineering Department for granting me access to high-resolution elevation data.

Conference presentations, short course attendance, and research development were supported by grants from the Department of Geology at the University of Georgia, International Association of Mathematical Geosciences, University of Georgia Office of the Vice President of Research, and Geological Society of America.

I am indebted to my family and friends for their continuous support and inspiration throughout my tenure in graduate school.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

PREFACE

The work presented hereafter has two components. The main component is contained within Chapters 2 – 5. Chapters 2 – 5 will be submitted for publication in the journal Earth Surface Processes and Landforms, and each chapter represents the sections of the manuscript (i.e., introduction, methodology, results, conclusion). Content and figures have been added in some areas for lengthened descriptions. The methodologies developed and results found in Chapters 2 – 5 are for a local study area. The secondary component, contained in chapter 6, demonstrates how the local model described in chapters 2 – 5 can be applied to a larger study area (area: 868 km^2). Chapter 7 is an overall summary of the two components.

While the two components are consistent with respect to scientific questions addressed and methodologies applied, the two components differ in data source, data resolution (i.e., spatial and temporal), study area scale, and types of data utilized. Additionally, the secondary component (Chapter 6) is content I generated while leading a research team in the NASA DEVELOP Applied Sciences Program at the University of Georgia from June to August (2015). Dr. Adam Milewski and I wrote a successful proposal to the NASA DEVELOP Program that was derived from the methodologies and results described in Chapters 3 and 4, respectively.

CHAPTER 2

INTRODUCTION

Karst topography can be found at any latitude and elevation around the world, with rock units potentially containing karst features covering approximately 20% of the Earth's land surface (Stokes *et al.,* 2010; Ford and Williams, 1989). Karst landscapes are characterized by three primary morphological features: input landforms that direct surface water underground (i.e., sinkholes), subsurface conduit systems (i.e., fractures and caves enlarged by solution), and discharge areas (i.e., springs) (Ford *et al*., 1988). Karst features, specifically sinkholes, present hazards and engineering challenges to residential, commercial, industrial, and agricultural infrastructure, serve as entry points for groundwater contaminants, and cause vertical deformation (i.e., subsidence) by transporting sediment underground (Hyatt and Jacobs, 1996; Waltham *et al.,* 2005; Galve *et al.,* 2009a; Newton, 1987).

Sinkhole formation is the result of complex interactions between hydrologic (e.g. flooding), geologic (e.g., overburden thickness), geomorphologic (e.g., elevation), anthropogenic (e.g. land use), climatic (e.g. precipitation), hydrogeologic (e.g. aquifer fluctuations), and other factors acting with fluctuating magnitudes over varying spatial and temporal scales. It is difficult to directly observe and quantify the influence of each factor responsible for sinkhole formation because the majority of the factors operate in the subsurface and over generally long time scales. Additionally, two or more of these

processes often operate in conjunction to form or enlarge an existing sinkhole (Ford *et al*., 1988). However, inferences about the controls on sinkhole formation can be made by measuring the relationships between sinkhole density and the spatial distribution of the factors that control sinkhole occurrence (Doctor and Doctor, 2012).

Literature Review

Although sinkhole formation is site-specific, relationships between sinkhole density and controlling factors can be determined in areas with historical data on sinkhole development and an abundance of ancillary datasets to accurately represent sinkhole formation controlling factors and mechanisms (Wilson and Beck, 1992). Panno *et al.* (2013) related sinkhole location and evolution to hydrogeologic (e.g., water table depth and storage coefficients), hydrologic (e.g., recharge rates), and geologic (e.g. bedrock topography) factors in the sinkhole plain of Illinois, USA. Hubbard (2001) analyzed sinkhole distribution in the Valley and Ridge Province, Virginia, USA, and correlated the highest sinkhole densities with lithology (e.g., bedding planes), geologic structures (e.g., fold and fault axes), and hydraulic gradients related to proximity to incised segments of rivers. Al-Kouri *et al.* (2013) found that sinkhole occurrence was most influenced by urban land use, fault distribution, and proximity to surface water features. In the Ebro Valley (Spain) evaporite karst setting, sinkhole susceptibility and hazard have been determined by quantifying the relationships between sinkhole type and distribution and different geomorphologic units, elevation, alluvium thickness, piezometric surface , land use, and electrical conductivity of the surficial aquifer (Galve *et al.,* 2009a,b; Gutiérrez *et al.,* 2007; Lamelas *et al.* 2008; Galve *at al.* 2008). Doctor and Doctor (2012) and Doctor

et al. (2008a,b) measured the influence of geologic (e.g., distance to fractures and fold axes) and hydrologic features (e.g., distance to streams and ponds) to sinkhole location Virginia, West Virginia, and Maryland, USA, karst regions. Gao and Alexander, Jr (2008) input bedrock type and overburden thickness to construct sinkhole probability maps in southeastern Minnesota and northwestern Iowa.

This study focuses on the mantled karst terrain of Dougherty County, Georgia, USA (Figure 2.1), an area with well-documented sinkhole development (Brook and Allison, 1986; Hyatt and Jacobs, 1996; Warner Gordon *et al.*, 2012; Hyatt *et al.*, 2001). Brook and Allison (1986) used topographic maps and 1:24,000 scale, color infrared images to identify sinkholes based on the presence of surface water features, vegetation and soil moisture patterns, and topographic expression. The mapped sinkhole distribution and color infrared images were used to map fractures and compare the mapped fracture orientations with regional trends. Hyatt and Jacobs (1996) found that flooding of the Flint River in 1994 triggered the formation of at least 312 sinkholes in and around the Albany area of northern Dougherty County; 88% of which formed within flooding limits. The mechanisms involved in the formation of the 312 sinkholes were liquefaction of the unconsolidated overburden into bedrock cavities and loss of buoyant support as flood waters receded. Hyatt and Jacobs (1996) noted that the sinkholes followed a linear pattern, which suggests that joints and fractures influence sinkhole distribution. Following Tihansky (1999), Warner Gordon *et al.* (2012) suggested that rapid fluctuations of the Upper Floridan aquifer and overburden removal (2.5 - 4.5m) for construction caused localized sinkhole formation in a municipal well field. In the covered karst region of Lowndes County, Georgia, Hyatt *et al.* (2001) suggested that several

factors had an influence on sinkhole locations, including elevation, soil type, overburden thickness, and potentiometric head levels.

Of the many approaches used to understand and/or predict sinkhole formation, Galve *et al.,* (2009b) found that nearest neighbor and sinkhole density methods perform better than other techniques when identifying areas of sinkhole susceptibility, but those methods do not include sinkhole formation explanatory variables. Thus, their ability to measure the influence of various factors on sinkhole development is limited (Doctor and Doctor, 2012). Geographically weighted regression (GWR) is a technique used to measure spatially varying relationships, such as the influence of controlling factors on sinkhole formation (ESRI 2012). GWR models a dependent variable by building a unique regression equation for individual points (e.g. sinkholes) and weighting the influence of each independent variable based on distance from the position of the dependent variable. GWR analysis results can be used to measure the overall fit of a model and quantify the degree of influence of each independent variable on a given dependent variable value.

Objectives

Previous studies have not quantified the relationships between sinkhole occurrence and a multitude of sinkhole formation factors in the study area. The goal of the present work was to evaluate the influence of controlling factors on sinkhole distribution in the mantled karst environment of southern Dougherty County, Georgia. The first objective was to produce sinkhole inventory maps for 10m resolution Digital Elevation Models (DEMs) from 1999 and 2011. Sinkholes that formed or enlarged between 1999 and 2011 were identified, hereafter referred to as temporal-difference (TD)

sinkholes, for a spatiotemporal analysis of sinkhole formation. The second objective was to create a high-resolution sinkhole inventory map from a 1m resolution LiDAR (Light Detection and Ranging) DEM. Spatial statistical techniques were utilized on these two sinkhole inventory datasets (TD and LiDAR) to fulfill the third objective of measuring the influence of controlling factors on sinkhole formation. Finally, sinkhole prediction maps were produced from spatial statistics results.

Study Area

The study area, covering 183 km^2, is located in southern Dougherty County, Georgia, and is part of the Dougherty Plain region of the Coastal Plains Physiographic Province in southwest Georgia (Figure 2.1). Precipitation averages 1270 mm/year but has high annual variability (Stewart *et al.,* 1999). Long-term precipitation patterns determined from a 12-month Standardized Precipitation Index show drought conditions occurred three separate times during the study period: 1998 – 2002, 2006 – 2008, and 2011 – 2013.

The karst topography is characterized as flat to gently sloping with land-surface altitudes ranging from 42 – 97 m within the study area (Stewart *et al.,* 1999). The lowest elevations are in the Flint River valley, and the land surface gains elevation as it slopes towards the Solution Escarpment along the eastern and southeastern portions of the study area (Figure 3.2). The karst topography of the Solution Escarpment, described by MacNeil (1947), differs from lower-lying areas due to greater overburden thicknesses (up to 45m) and relatively infrequent sinkhole formation (Hicks *et al.,* 1987).

Land use is primarily forested (43.7%), cultivated crops (23.4%), and grassland (15.6%) (Homer *et al.*, 2015) (Figure 3.9). Central pivot irrigation systems are used for many cultivated crops (e.g. cotton and peanuts), and this agricultural practice may enhance sinkhole formation by increasing infiltration processes and accelerating erosion of incohesive unconsolidated materials into preexisting, solution-enlarged subsurface cavities (Ford and Williams, 1989). Minor roads are located in developed locations, but several major roads and highways cross the study area. Proximity to roads was a factor considered in the sinkhole formation analysis, as human activities that alter the physical environment (e.g., construction and groundwater extraction) may act as mechanisms that trigger sinkhole formation (Rose et al., 2004; Galve *et al.,* 2009a; Waltham *et al.,* 2005; White *et al.,* 1986).

The surface hydrologic features considered include streams, ponds, and wetlands (Figure 3.7). Internal drainage due to active limestone dissolution has limited surface drainage in the study area (Hicks *et al.*, 1987; Warner Gordon *et al.,* 2012). The main surface drainage feature is the Flint River, which forms a terraced valley with a floodplain ranging from 1 – 4 km wide in the study area. The primary tributary to the Flint River in the study area is Dry Creek, an intermittent stream flowing from the highlands of the Solution Escarpment. Man-made ponds and intermittent ponds, which form from surface runoff depositing clay and other fine-grained material into paleo-sinkholes to create an impermeable base layer, occur throughout the study area (Ford and Williams, 1989). The majority of floodplains of the Flint River and Dry Creek are classified as wetlands (USFWS, 2002). Other wetlands are primarily east of the Flint River and north of Dry Creek.

The geology of southern Dougherty County is characterized by alternating units of Cretaceous to Quaternary age sandstone, dolomite, limestone, shale, and siltstone (Hicks *et al.,* 1987). In this study, the Ocala Limestone was the only geologic unit considered due to its influence on sinkhole formation. The Ocala Limestone of late Eocene age underlies the unconsolidated overburden, is densely fractured in areas (up to 5.5 fractures / km^2), and has well-developed secondary permeability along the fractures, joints, and bedding planes due to solution enlargement (Brook and Allison, 1986; Hicks *et al.,* 1987). The unconsolidated overburden (i.e., soil) ranges from 0 – 40 m thick in the study area (Figure 3.6) and is primarily made up of fine, well-drained to poorly drained loamy sands, sandy loams, clays, and weathered bedrock material. The thinnest layer of overburden thickness is along the Flint River where the Ocala Limestone outcrops due to erosional downcutting. The unconsolidated overburden is thinner west of the Flint River (< 25m) and progressively thickens toward the southeast of the study area to 40m of unconsolidated overburden (Brook and Allison, 1986; Parker and Hawman, 2012).

The hydrogeology is comprised of the surficial aquifer and the Upper Floridan aquifer (UFA). The surficial aquifer has an intermittent areal extent and, where it exists, is classified as a perched water table on shallow clay lenses (Torak *et al.,* 1993). The Ocala Limestone comprises the upper hydrostratigraphic unit of the UFA. In the northern region of the study area, transmissivity values derived from diffusivity analyses and multiwell aquifer performance tests range from 65 to 26,300 m^2 per day (Torak and Painter, 2006). Specific capacities, storage coefficients, and transmissivity values vary substantially due to limestone dissolution and fractured bedrock (Brook and Sun, 1982; Hicks et al., 1987). Aquifer levels exhibit annual fluctuations up to 14m (Figure 3.8), but

are generally highest throughout the winter season into spring months and lowest during summer and fall months due to precipitation patterns and increased groundwater extraction for agricultural purposes.

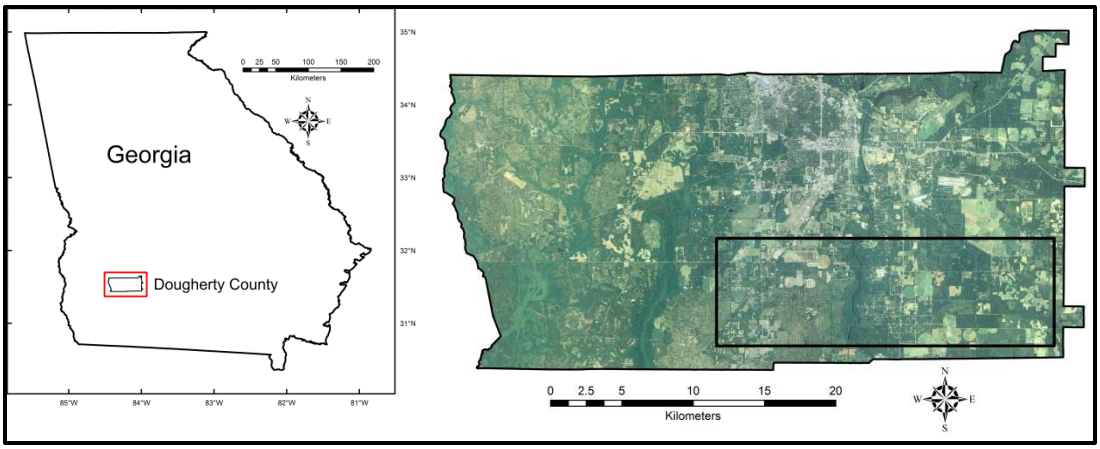

Figure 2.1: Dougherty County is located in southwest Georgia. The 183 km^2 study area is outlined in southeast Dougherty County.

CHAPTER 3

METHODOLOGY

The integrated approach towards identifying and analyzing sinkhole formation involved three steps. First, we identified and mapped sinkholes using three DEMs with varying horizontal resolutions: 1999 (10 m), 2011 (10 m), and 2011 LiDAR (1 m). Two sinkhole inventory datasets were extracted from the sinkhole mapping procedure: 1) temporal-difference (TD) and 2) LiDAR. The TD map represents sinkholes that formed or enlarged between 1999 and 2011 and were identified by comparing the 1999 (10 m) and 2011 (10 m) sinkhole mapping results. A higher-resolution sinkhole inventory was derived from the 1 m LiDAR DEM. Filter mechanisms and validation measures were applied to improve the accuracy of the sinkhole maps. Secondly, a sinkhole formation model was developed for southern Dougherty County by measuring the relationships between the previously mapped sinkhole distributions and a variety of sinkhole formation factors. Ordinary least squares (OLS) and geographically weighted regression (GWR) spatial statistical techniques were utilized to identify overall model performance and determine the most influential factors responsible for the TD and LiDAR sinkhole distributions. Thirdly, the GWR results were tested for their predictive capability by using the GWR sinkhole density estimates in the study area in a deterministic interpolator (e.g. Inverse Distance Weighting). The resulting TD and LiDAR predictive maps were compared to the TD and LiDAR measured maps.

Sinkhole Mapping

A complete sinkhole inventory is essential for accurately analyzing the factors influencing sinkhole development (Gutiérrez *et al.*, 2011). However, it is time and money-intensive to map sinkholes using traditional methods (e.g., aerial imagery interpretation, field observation, and topographic maps), but the availability of DEM products have allowed for improved efficiency and accuracy of sinkhole inventories, especially at larger scales. Additionally, this work utilizes the availability of current and historical DEMs, which allows for a spatiotemporal analysis of sinkhole formation. We present a sinkhole mapping methodology adapted from Rahimi and Alexander, Jr. (2013), Shaw Faulkner *et al.* (2013), and Doctor and Young (2013) to produce reliable sinkhole inventories from which sinkhole density measurements were derived (Figure 3.1). Evaluations of the adopted sinkhole mapping methodology show that 80 – 90% of sinkholes are correctly identified (Rahimi and Alexander, Jr., 2013). Sinkhole mapping and subsequent analyses were performed within ArcGIS 10.1 (ESRI 2012).

DEMs from 1999 and 2011 were utilized for sinkhole inventory production. The 1999 DEM (10m horizontal resolution) was obtained from the U.S. Geological Survey National Elevation Dataset (Gesch *et al.*, 2002). The 2011 LiDAR DEM (1m horizontal resolution) was gathered from the geographic information system (GIS) division of the City of Albany and Dougherty County Engineering Department. The 2011 LiDAR DEM was resampled to a 10m horizontal resolution DEM using a cubic interpolation algorithm (Kidner *et al.*, 1999). The stream and road buffer polygon layers, which were used as filtering mechanisms to minimize falsely identified sinkholes (i.e., false positives), were derived from the USGS National Hydrography Dataset

(http://datagateway.nrcs.usda.gov) and the Georgia GIS Clearinghouse (https://data.georgiaspatial.org), respectively. High-resolution aerial imagery (1-2m) from 2005-2013 was used to validate sinkhole mapping results. These data were gathered by the US Department of Agriculture (USDA) National Agricultural Imagery Program (NAIP) available at http://datagateway.nrcs.usda.gov. Additionally, historical Google Earth imagery for 1999, provided by the U.S. Geological Survey, and historical sinkhole maps from Brook and Allison (1986), Hyatt and Jacobs (1996), and Warner Gordon *et al.* (2012) were used during the sinkhole inventory validation procedure.

The first step to produce a sinkhole inventory map using GIS is to fill DEM depressions to their spill level, or height where a drainage direction can be defined (ESRI 2012). The original DEMs were then subtracted from the filled DEMs. These differenced data effectively identify each depression (i.e., potential sinkhole) in the original DEM and provide important geometric characteristics (e.g., depth) of the depressions (Doctor and Young, 2013).

Filter mechanisms were applied to improve the accuracy of the sinkhole inventories by removing falsely identified sinkholes initially recognized from the differenced data. Geometric characteristics, including depth, eccentricity, and area, were calculated and attributed to each sinkhole polygon. The depth filter was based on the vertical accuracies of the DEMs, which are 1.55 m root-mean-square error (RMSE) for the 1999 DEM and 20cm RMSE for the 2011 LiDAR DEM. The 2011 10 m DEM was conservatively assigned a 1.55 m RMSE vertical accuracy to remain consistent with the 1999 sinkhole inventory production. Application of the depth filter results in a 95% confidence level that an identified depression is a true depression in the DEM (Doctor

and Young, 2013). The second geometric filter applied to the sinkhole inventory datasets was based on the eccentricity of an ellipse (e), which is calculated as:

$$e = \sqrt{1 - \frac{b^2}{a^2}} \qquad \textit{Equation 1}$$

where *a* and *b* are one-half of the depression's major and minor axes, respectively. The geometry of the depression diverts farther from a perfect circle as eccentricity values approach 1 and becomes more compressed due to greater differences between the major and minor axes lengths. Based on aerial imagery and visual analysis of the sinkhole datasets, an eccentricity threshold of 0.96 was applied to remove elongated depressions that were unlikely to be true sinkholes. Lastly, an area filter mechanism was applied to the LiDAR sinkhole dataset. Hyatt and Jacobs (1996) measured 53 sinkholes immediately north of the study area and found they had an average area of 11.48m^2. Due to the small sample size of the those results and high horizontal resolution (1m) of the LiDAR DEM, we chose an area threshold of 3m^2 to eliminate any falsely identified sinkholes that could be the result of horizontal accuracy errors.

Filtering mechanisms related to natural and anthropogenic artifacts must be applied to sinkhole inventory datasets to exclude falsely identified sinkholes that are, for example, a result of misinterpretation of shallow or intermittent stream channel features or road construction (e.g., ditches) (Shaw Faulkner *et al.,* 2013). A buffered polygon file was created that represents the area 5m from either side of the middle of the stream channel and its intermittent tributaries to address natural artifacts related to intermittent stream channels in the study area (i.e., Dry Creek). Similarly, a road buffer was generated that represents the area 15 m from either side of the central road line to eliminate falsely identified sinkholes in constructed road ditches. Proximity filters were

applied that eliminated sinkholes with greater than 50% of their area within the stream and road buffers.

A two-fold manual validation process was implemented to verify the accuracy of the sinkhole inventories and eliminate any falsely identified sinkholes that remained after the filter mechanisms were applied. First, georeferenced, high-resolution (1-2 m) aerial imagery was used to verify sinkhole locations. This process was primarily focused on urban areas and anthropogenic structures. For example, sinkholes that fell within 3m of building footprints were eliminated. Secondly, previous sinkhole maps from Brook and Allison (1986), Hyatt and Jacobs (1996), and Warner Gordon *et al.* (2012) were compared with the results of the applied sinkhole mapping methodology. One issue with this approach is the local land use practice of filling in sinkholes, especially in urban areas or near roads, thus excluding them from being identified with our mapping procedure. However, large-scale sinkholes and areas of high sinkhole density identified by previous studies matched well with our sinkhole mapping results. Despite the thorough manual validation procedures, sinkhole mapping results should be considered as minimum sinkhole inventories due to DEM resolutions and methodology constraints (Galve *et al.,* 2008; Rahimi and Alexander, Jr., 2013).

The results of the 1999 (10 m) sinkhole map were subtracted from the 2011 (10 m) sinkhole map to produce the temporal-difference (TD) sinkhole dataset, which represents sinkholes that formed or were enlarged between 1999 and 2011 (Figure 3.2). The two sinkhole polygon datasets (i.e., TD and LiDAR) were prepared for further analyses upon completion of the validation procedure.

Cluster analyses were performed on the sinkhole datasets to quantify spatial distribution patterns and provide empirical distance inputs for GWR bandwidths. Figure 3.3 shows the sinkhole distribution pattern results for the TD and LiDAR datasets. The bandwidth applied to the GWR analyses was the distance of maximum clustering derived from the cluster analysis and differed from the bandwidth used to measure sinkhole density to avoid presenting dependence into the GWR model (Doctor and Doctor, 2012).

Sinkhole density was measured at each sinkhole point location using GIS. A 1 km radius was used to calculate sinkhole densities for both sinkhole datasets to allow for proper spatial variance of density values for the purposes of GWR calculations. Figures 3.4 and 3.5 show the TD and LiDAR sinkhole density results, respectively, expressed as the number of sinkholes per km^2. The sinkhole density value was spatially assigned to each sinkhole point as the dependent variable to be modeled in the OLS and GWR analysis.

Sinkhole Formation Mechanisms

Five separate categories of independent variables (i.e., predictor or explanatory variables) were implemented into the spatial statistical models that represent the processes deemed to be governing the spatial distribution of sinkhole formation in the study area. The processes (e.g., geologic, hydrologic), specific variables, data sources, and method of production are presented in Table 3.1. GIS was used to import and georeference variable layers that were acquired from aforementioned sources (e.g., hydrography, wetlands, roads, etc.). Interpolated variable layers (e.g., aquifer fluctuations and overburden thickness) were produced within GIS. Independent variables that

represent pedologic, geochemical, and climatic processes could not be included in our sinkhole formation model due to paucity of data, insufficient data form (i.e., low horizontal resolution), and/or inability to properly represent the sinkhole formation mechanism (e.g., individual high-precipitation events). However, qualitative analyses can be made regarding the unrepresented processes and is discussed later. A description of each independent variable is included below:

- *Proximity to fracture (geologic)*: Distance to nearest bedrock fracture (i.e., joint or lineament). Bedrock fracture data from Brook and Allison (1986) were georeferenced to the study area and used to find the distance from each sinkhole to the closest fracture (Figure 3.6). As noted by Brook and Allison (1986) and Hyatt and Jacobs (1996), sinkholes tend to follow fracture-controlled linear trends in Dougherty County. Fractures allow for high secondary permeability, which increases dissolution potential of the Ocala Limestone and creates subsurface voids for overburden material to be transported and deposited; thus forming a sinkhole.
- *Overburden thickness (geologic)*: Residuum thickness (i.e., depth to bedrock) overlying the Ocala Limestone (Figure 3.6). Borehole and cross-section point data (n = 48) were gathered from Hicks *et al.,* 1987; Torak *et al.,* 1993; Torak and Painter, 2006; McSwain, 1998; Warner, 1997; McFadden and Perriello, 1983; Clarke *et al.,* 1984. This variable layer was generated using the Empirical Bayesian Kriging geostatistical interpolator (RMS = 2.42). This method constructs local models on subsets of the data, thus accounting for some of the

local heterogeneity of bedrock topography and residuum thickness in the mantled karst terrain.

- *Proximity to stream (hydrologic)*: Distance to nearest perennial (Flint River) and intermittent (Dry Creek) stream channels (Figure 3.7). This variable reflects areas prone to flooding, which can initiate sinkhole formation by saturating and liquefying the unconsolidated overburden, reducing soil stability, and accelerating soil transportation into bedrock cavities (Hyatt and Jacobs, 1996).
- *Proximity to wetland (hydrologic)*: Distance to the nearest swamps and marshes identified from the USFWS National Wetland Inventory (Figure 3.7). Wetlands are thought to influence sinkhole formation in the study area due to soil moisture conditions and infiltration processes that may lead to increased dissolution rates, bedrock weathering, and increased transport of residuum materials into subsurface voids.
- *Proximity to pond (hydrologic)*: Distance to the nearest pond (Figure 3.7). Leakage from unlined ponds supports subsurface erosion and dissolution processes (Galve *et al.,* 2008). Ponds have been shown to drain due to sinkhole formation within or in the immediate vicinity of ponds in the study area (J. Stolze, Albany Utilities, personal communication, June 26, 2015).
- *Aquifer fluctuations (hydrogeologic)*: Average UFA oscillations from 1999 – 2011 using USGS National Water Information System well data (Figure 3.8). Warner Gordon *et al.* (2012) has shown that sinkhole formation correlates with rapid water level fluctuations in the study area. The maximum and minimum well level differences for each year from 1999 to 2011 were averaged (n = 12). The

georeferenced point data were input into a kriging interpolator to produce a smooth surface (RMS = 3.22).

- *Land use (anthropogenic)*: Land use and land cover type ascribed by the USGS National Land Cover Database (Figure 3.9). Human activities that modify the physical environment (e.g., construction, groundwater extraction, irrigation, etc.) have been shown to influence sinkhole distribution (Rose et al., 2004; Galve *et al.*, 2009a; Waltham *et al.*, 2005; White *et al.*, 1986). This variable was not included in GWR analysis due to the data type (categorical).
- *Proximity to road (anthropogenic)*: Distance to nearest road identified in the U.S. Census Bureau roads and highways inventory (Figure 3.9). Road excavations decrease overburden thickness and change local drainage directions and infiltration rates due to ditch construction.
- *Elevation (geomorphologic)*: Ground surface altitude in meters a.s.l. derived from the LiDAR DEM representing the geomorphology of the study area (Figure 3.2). This variable was not included in GWR analysis due to non-stationarity, which is most likely the result of the gently sloping land surface where elevation values have low variance.
- *Aspect (geomorphologic)*: Direction of the steepest land surface slope gradient obtained from the LiDAR DEM. This layer represents topographic controls on dominant local drainage directions which may influence subsurface infiltration and erosion processes. Similar to elevation, this variable was not included in GWR analysis due to non-stationarity.

Spatial Statistics

Following the processing of the aforementioned variables, ordinary least squares (OLS) regression was used to evaluate the performance of each sinkhole formation model (e.g., TD and LiDAR) and verify the models do not violate common regression model assumptions through diagnostic statistics assessments. OLS applies a global linear regression model for the entire study area to measure the overall fit of the data and produce predictions for a dependent variable (i.e., sinkhole density) based on relationships to assigned independent variables (i.e., geologic, hydrologic, anthropogenic, etc.). The diagnostic statistics showed that neither sinkhole formation models violated these important assumptions: the modeled relationships were consistent (Koenker (BP) statistic), model predictions were normally distributed and not biased (Jarque-Bera statistic), and redundancy among explanatory variables was not an issue (low variance inflation factors < 7.5).

OLS assumes variable relationships are constant in space, which prohibits regression coefficients to vary over space. Due to the spatial heterogeneity of sinkhole density and individual sinkhole formation factors, a local spatial statistics technique that allows regression coefficients to vary, such as GWR (Brunsdon *et al.*, 1998), is better suited for quantifying the influence of sinkhole formation controls and producing more accurate predictions of sinkhole density. GWR provides a direct method of testing hypotheses that are the subject of spatial varying relationships by computing a unique regression equation for each point in a dataset using an assigned kernel type, bandwidth method, and optional weighting factors for individual features.

GWR investigations are influenced by the size of the study area and the regression analysis bandwidth (Doctor and Doctor, 2012). The selected bandwidth defines the area around each sinkhole that the regression equation considers; that is, what value(s) for each sinkhole formation factor (e.g. independent variable) to consider for modeling sinkhole density. In this study, the GWR bandwidth selection is a fixed distance based on cluster measurements. Utilizing a fixed bandwidth has been shown to produce useful model results when the regression model primarily considers proximity data (Doctor and Doctor, 2012). However, sinkholes located in low sinkhole density areas did not receive a regression prediction due to low data availability.

In addition to measuring the influence of controlling factors on sinkhole formation, GWR sinkhole density prediction results are input into geostatistical interpolators to test the ability of GWR to predict sinkhole density. The TD and LiDAR measured sinkhole density results were compared to the GWR-derived predicted sinkhole density results to identify trends and limitations of GWR as a predictive tool.

Table 3.1: Summary of variables used in the OLS global sinkhole formation model. The variables used in the GWR local model are indicated in bold.

	Variable	Source	Production Method
Geologic	**Proximity to fracture**	Brook and Allison, 1986	ArcMap near tool
	Overburden thickness	USGS borehole data	Empirical Bayesian Kriging interpolation
Hydrologic	**Proximity to stream**	USGS National Hydrography Dataset	ArcMap near tool
	Proximity to wetland	USFWS National Wetland Inventory	ArcMap near tool
	Proximity to pond	USGS National Hydrography Dataset	ArcMap near tool
Hydrogeologic	**Aquifer fluctuations**	USGS National Water Information System groundwater data	Inverse Distance Weighting interpolation
Anthropogenic	Land use	USGS National Land Cover Dataset	-
	Proximity to road	U.S. Census Bureau	ArcMap near tool
Geomorphologic	Elevation	LiDAR and USGS NED	-
	Aspect	LiDAR and USGS NED	ArcMap aspect tool

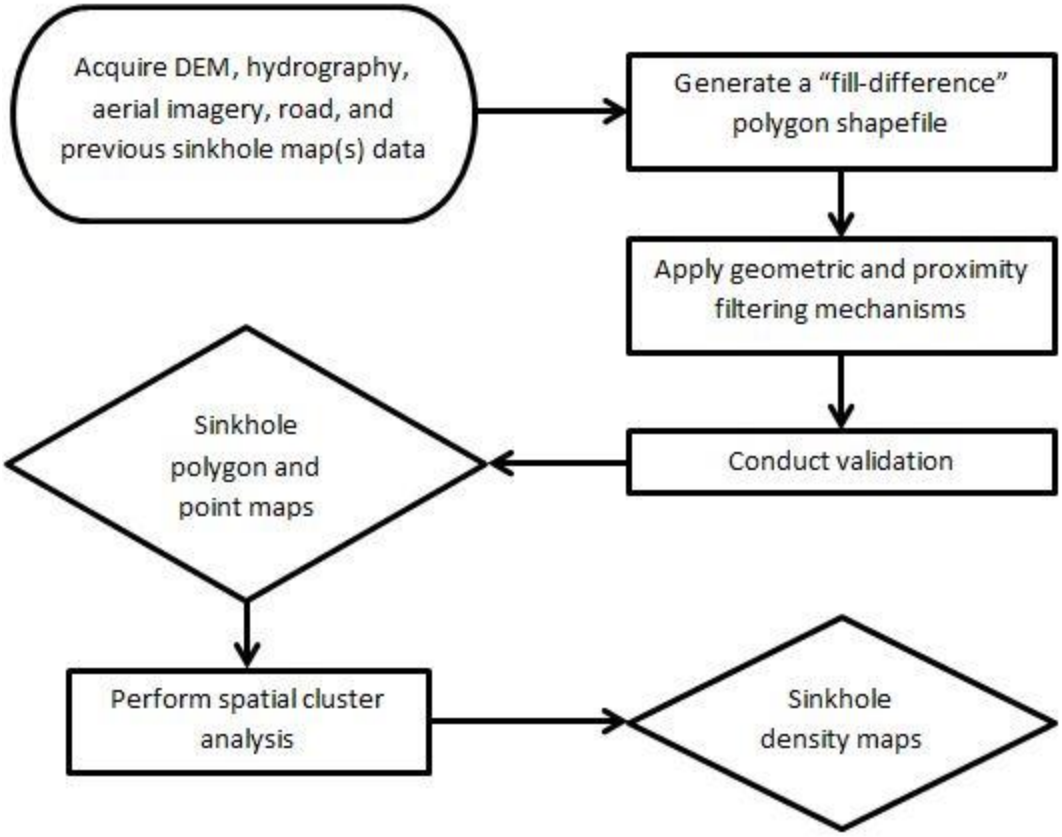

Figure 3.1: Generalized flow chart delineating the methodology to produce accurate sinkhole inventory maps and sinkhole density measurements.

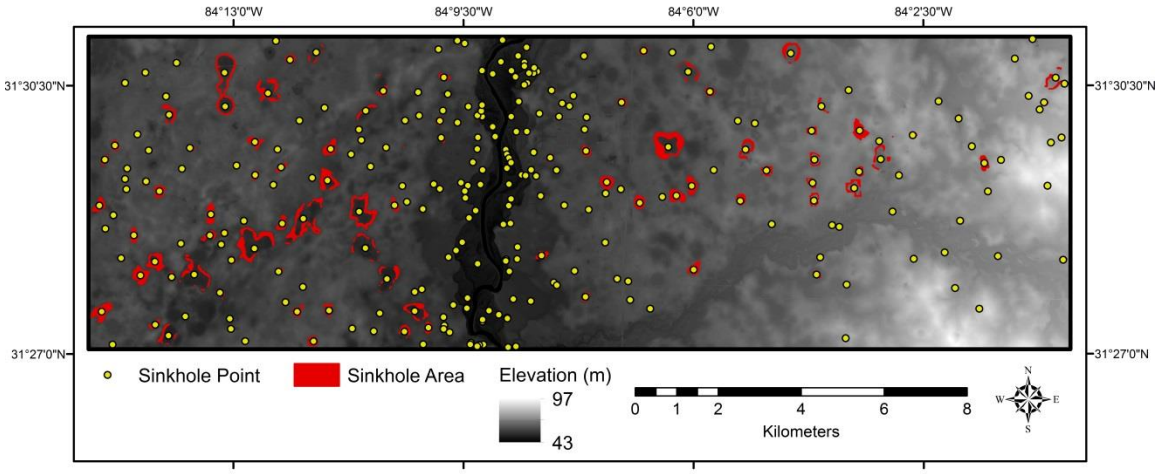

Figure 3.2: Map showing the results of the newly formed and enlarged sinkholes from the temporal-difference (TD) analysis.

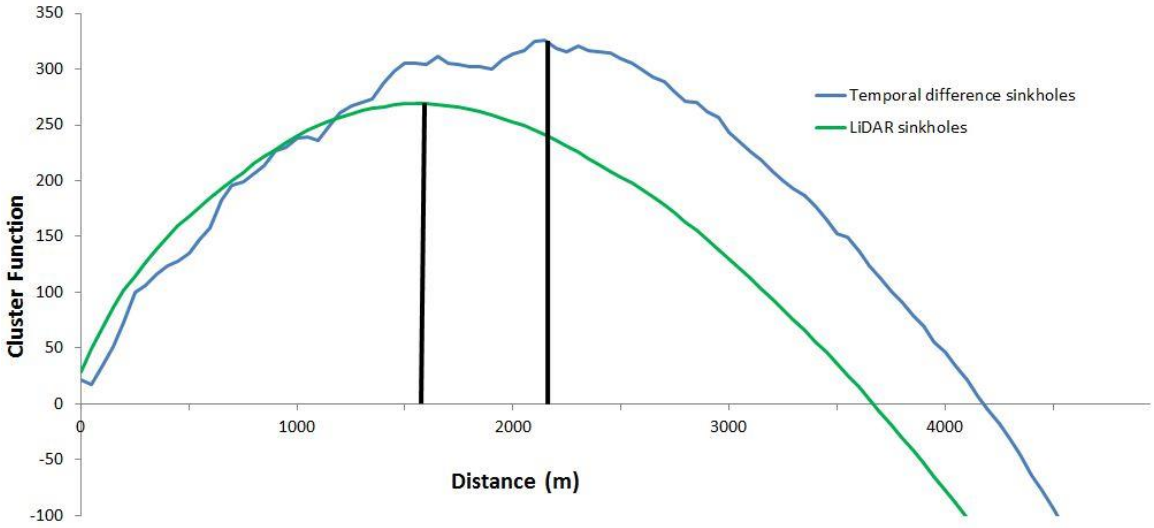

Figure 3.3: Multi-distance spatial cluster analysis plot specifying sinkhole distribution patterns for the mapped TD and LiDAR sinkholes. The vertical lines indicate the distance of maximum clustering for the two sinkhole datasets.

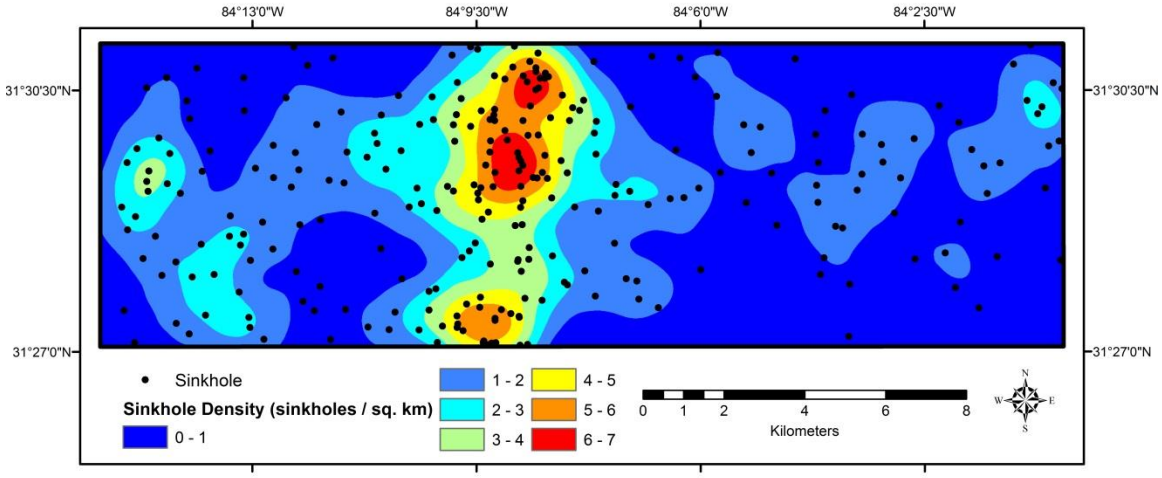

Figure 3.4: Sinkhole density map derived from the temporal-difference (TD) spatiotemporal analysis.

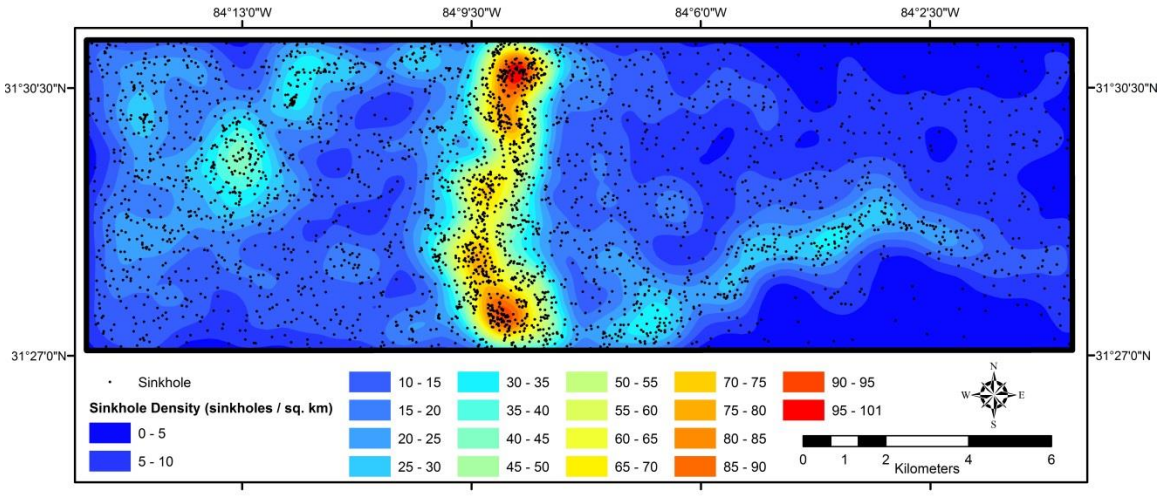

Figure 3.5: Sinkhole density map derived from the high-resolution LiDAR analysis.

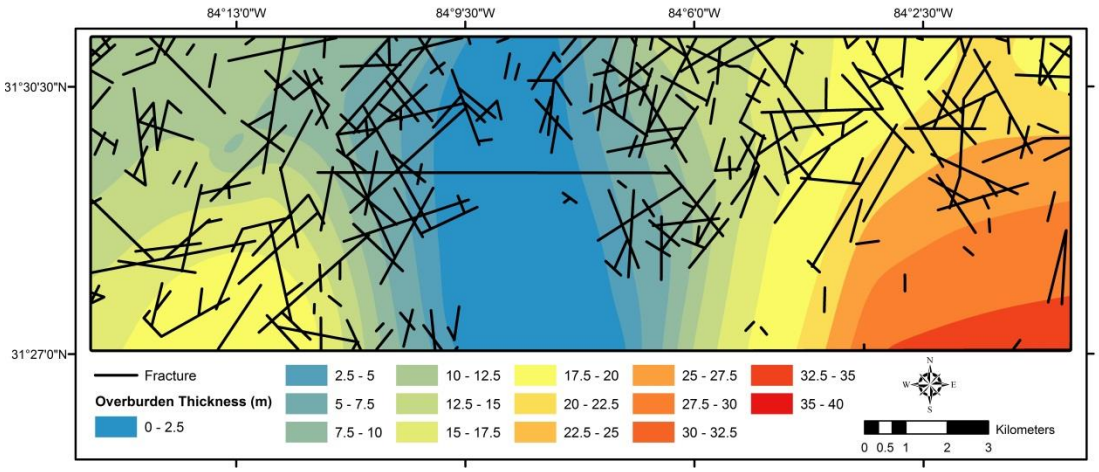

Figure 3.6: Geologic independent variables: proximity to bedrock fractures and overburden thickness (i.e., depth to bedrock).

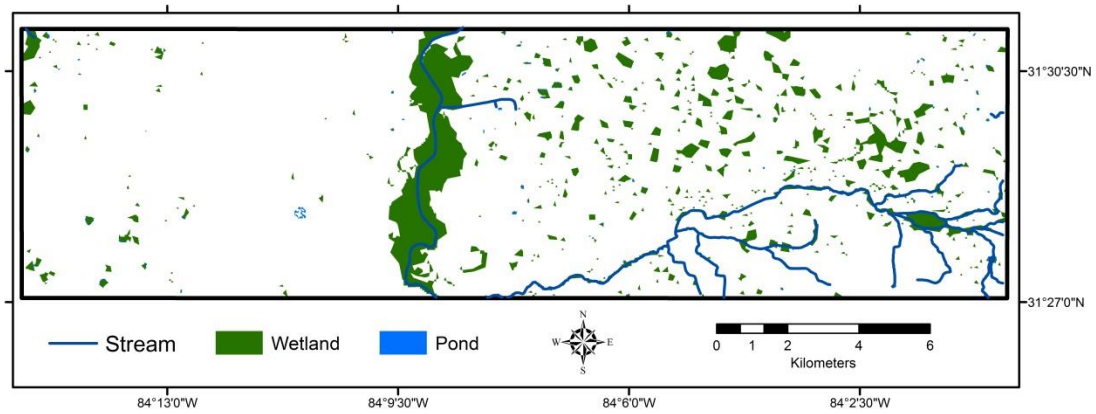

Figure 3.7: Hydrologic independent variables: proximity to streams, ponds, and wetlands.

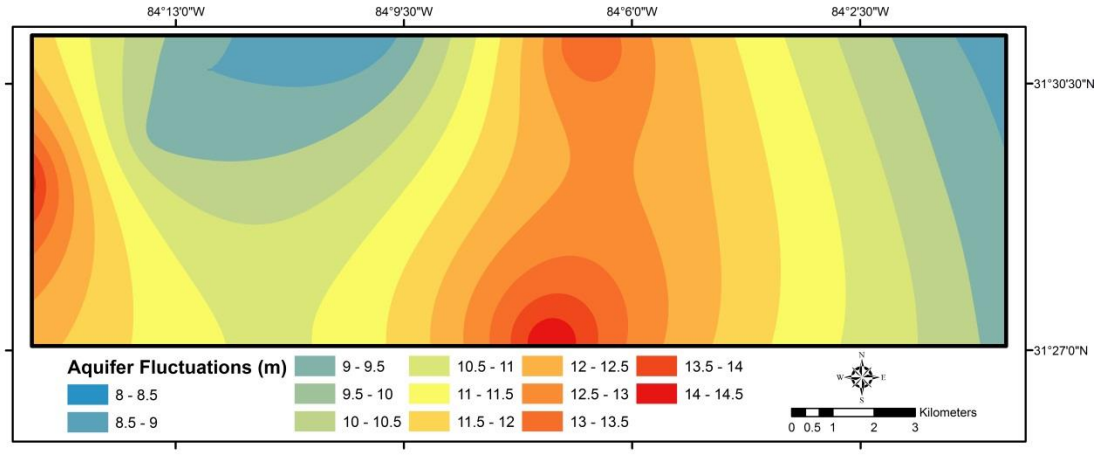

Figure 3.8: Hydrogeologic independent variable: Upper Floridan Aquifer fluctuations.

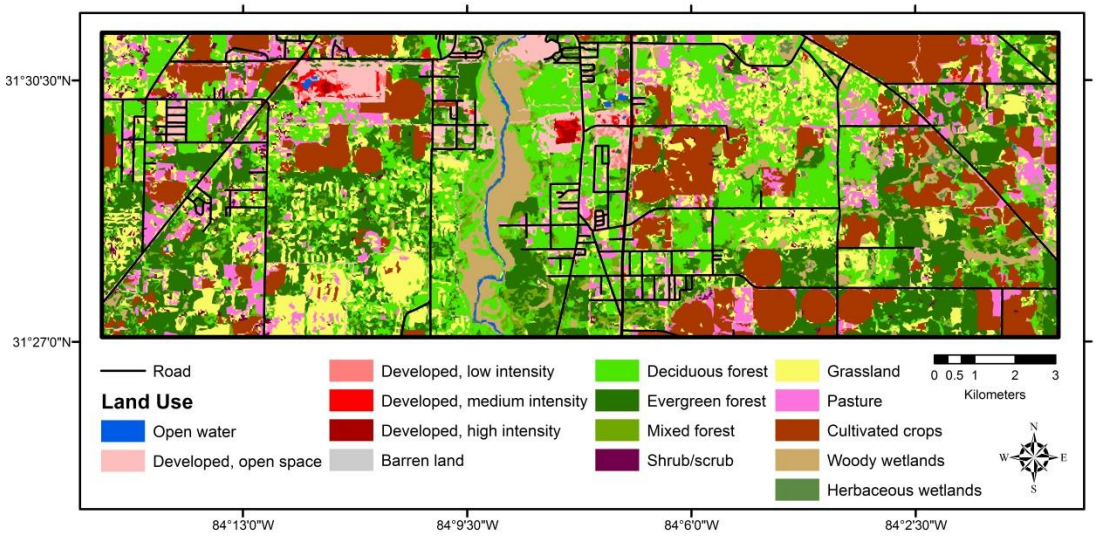

Figure 3.9: Anthropogenic independent variables: land use and proximity to roads.

CHAPTER 4

RESULTS AND DISCUSSION

Sinkhole Mapping and Spatiotemporal Evolution

The initial number of identified depressions from the three differenced sinkhole polygon layers were 1,273 (1999 10 m), 7,793 (2011 10 m), and 631,243 (LiDAR 1 m). Table 4.1 summarizes the results of the geometric and proximity filtering mechanisms applied to the initial sinkhole datasets. The depth filter mechanism allowed for a high percentage of possible false-positive depressions (78-86%) to be removed. 275 and 3,412 sinkholes were identified in the final TD and LiDAR datasets, respectively (figure 3.4 and 3.5).

Cluster analysis (i.e., nearest neighbor) results are summarized in table 4.2. Both datasets show statistically significant clustering at shorter distances and statistically significant dispersion at longer distances. Considering all distances, the TD sinkholes exhibit a randomly distributed pattern ($p = 0.86$, $z = 0.18$), and the LiDAR sinkholes showed a significantly clustered distribution ($p < 0.00$, $z = -6.27$). The random distribution of the TD sinkhole dataset is influenced by a smaller sample size in a large study area.

Sinkholes were classified into three categories based on area (measured as an ellipse): small = < 250 m^2; medium = 250 - 1,000 m^2; and large = > 1,000 m^2. The average sinkhole area was 10,208 m^2 and 6,947 m^2 for the TD and LiDAR datasets,

respectively. The spatial resolution differences largely influence the dissimilarity in average sinkhole size between the two datasets. For example, the smallest sinkhole that could be detected in the TD dataset was ~100m due to the 10m resolution DEMs used to derive that dataset. Additionally, the effect of the broad distribution of sinkhole areas in Dougherty County (Brook and Allison, 1986; Hyatt and Jacobs, 1996) was taken into account when classifying sinkholes. Large sinkholes (> 5,000 m^2) were excluded during classification due to their impact on average area despite their relatively low abundance. Instead, the average sinkhole area and standard deviation for sinkholes less than 1,000 m^2 were calculated and used for classification purposes (TD: 266 m^2 avg., 234 m^2 SD; LiDAR: 100 m^2 avg, 179 m^2 SD).

The spatiotemporal TD mapping results show small sinkholes (n = 151) formed within close proximity to streams in areas of shallow to intermediate overburden thickness. Medium sinkholes (n = 54) were mostly randomly dispersed across the study area, though several formed along the margins of the Flint River floodplain and in areas of high UFA fluctuations. No large sinkholes formed during the study period (e.g., 1999 – 2011), but pre-existing large sinkholes enlarged during the study period (n = 70) due to erosional processes.

The spatially-detailed LiDAR-derived sinkhole inventory exhibited similar spatial patterns to the TD sinkhole dataset. Small sinkholes (n = 2,394) cluster near streams in areas of shallow to intermediate overburden thickness primarily west of the Flint River. Medium sinkholes (n = 312) are found along floodplain margins in areas of intermediate overburden thickness. Large sinkholes (n = 706) were mainly distributed in areas

characterized by intermediate to deep overburden thickness. Large sinkholes were not present within floodplains.

Sinkhole Formation Model and Individual Factor Influence

Firstly, the GWR sinkhole formation models were assessed for overall goodness of fit between observed and predicted sinkhole density values using adjusted coefficient of determination values (adj. R^2). 94% and 92% of sinkhole density variability were explained in the TD and LiDAR sinkhole datasets, respectively (Table 4.3). The improvement from the global OLS to local GWR results reflects the capability of GWR to model variation inherent in the explanatory variables.

Once the overall model results were evaluated, the factors with the strongest explanatory capabilities for sinkhole distribution were identified by GWR p-value results (Table 4.4). Overburden thickness, aquifer fluctuations, and proximity to fractures, streams, and wetlands exhibited statistically significant explanatory power for both sinkhole datasets. Distance to roads was statistically significant for only the LiDAR-derived sinkhole dataset. Distance to ponds proved to not have a strong influence on sinkhole distribution. These results are consistent with previous studies in the area (Brook and Allison, 1986; Warner Gordon *et al.*, 2012; Hyatt and Jacobs, 1996) and are similar to results from other karst areas (Table 4.5).

The addition of explanatory variables related to pedologic (e.g., clay content of soil), geochemical (e.g., calcium carbonate saturation), and climatic (e.g., precipitation) factors would most likely improve the overall model results but could not be explicitly included in the sinkhole formation models due to data paucity or model criteria (i.e., data

type or non-stationarity). Regardless, data were analyzed to contribute to the understanding of sinkhole formation in Dougherty County. Figure 4.3 shows a time-series of Flint River discharge, daily precipitation, and groundwater levels. High precipitation events and flooding accelerate sinkhole formation through liquefaction processes (Hyatt and Jacobs, 1996). Sharp decreases in groundwater levels, a likely result of groundwater extraction, contribute to sinkhole formation by reducing groundwater's hydrostatic support of overlying unconsolidated materials, thus allowing for accelerated sediment transport into subsurface voids. Figure 4.4 and 4.5 depict rainfall and drought conditions, respectively, that occurred in the study area throughout the study period. PRISM (Parameter-elevation Relationships on Independent Slopes Model) annual precipitation amounts from each year during the study period were averaged to produce figure 4.4. Precipitation varies only slightly across the study area. Thus, no relationship was identified between the spatial distribution of rainfall and sinkhole development. Wet and dry periods were identified by calculating the Standardized Precipitation Index (1998 – 2002, 2006 – 2008, and 2011 – 2013) for the study area. Similar to aquifer fluctuations, oscillations between wet and dry periods may have also contributed to sinkhole formation due to variance in hydrostatic support.

Geostatistical Prediction Interpolation

Figures 4.1 and 4.2 compare the measured and predicted sinkhole density maps for the TD and LiDAR datasets, respectively. The results show general similarity, where areas of high and low sinkhole density compare well as a first-order estimate. The TD sinkhole density estimates over-predict in areas of high density, while the LiDAR-

measured high sinkhole density areas are under-predicted, specifically in areas with > 80 sinkholes/km^2. Areas that had poor predictions are attributed to the spatial autocorrelation of regression residuals in areas of high spatial variability in sinkhole density measurements. These results show that, although robust, the GWR technique may not be the best method for predicting sinkhole distribution patterns. Rather, GWR is best suited for delivering a quantitative understanding of independent variable influences on sinkhole distribution patterns.

Table 4.1: Results of sinkhole mapping filter mechanisms. Values represent number of sinkhole points in the datasets.

Filter Mechanism	**1999 (10m)**	**2011 (10m)**	**LiDAR (1m)**
Depth	183	1,583	134,571
Eccentricity	179	1,493	119,159
Area	-	-	21,294
Stream buffer	173	1,255	17,967
Anthropogenic buffer	160	497	3,412
Final number of sinkholes	160	497	3,412

Table 4.2: Nearest neighbor and spatial cluster analysis results.

	Temporal Difference	**LiDAR**
Number of observations	275	3,412
Nearest neighbor classification	Random	Clustered
Nearest neighbor z-score	0.18	-6.27
Distance of maximum clustering	2,350 m	1,780 m
Distance to dispersed pattern	3,880 m	3,560 m

Table 4.3: Results summary of the OLS and GWR models. AICc values account for model complexity and goodness of fit. The results from the sinkhole formation models with the lowest AICc values were used to produce the prediction maps.

	Temporal Difference	**LiDAR**
OLS Adj. R^2	0.69	0.69
OLS AICc	795	27,310
GWR Adj. R^2	0.94	0.92
GWR AICc	174	5,125

Table 4.4: GWR independent variable p-value results for the sinkhole datasets. Significant at *** 0.1% level; ** 1% level; * 5% level.

Explanatory Variable	Temporal Difference	LiDAR
Proximity to fracture	0.00***	0.00**
Overburden thickness	0.00**	0.00**
Proximity to stream	0.02*	0.00**
Proximity to wetland	0.03*	0.00**
Proximity to pond	0.06	0.07
Aquifer fluctuations	0.03*	0.01*
Proximity to road	0.06	0.04*

Table 4.5: Results summary of influential sinkhole formation variables in other karst regions. Doctor and Doctor, 2012 (a): West Virginia; Galve *et al.,* 2008 (b): Spain; Doctor *et al.,* 2008 (c): Maryland; Lamelas *et al.,* 2008 (d): Spain; this study (e): Georgia.

Category	**Influential Variables**
Geologic	Lithologic gradient (b), overburden thickness (b, e), Quaternary deposit thickness >30m (d), low percent of Quaternary impermeable layers (d), and distance to nearest: fault (a,c), fold axes (a,c), fracture (a,e)
Hydrologic	Distance to nearest: stream (a,e), spring (a), pond (c), endorheic/wetland area (d,e)
Anthropogenic	Irrigation location, distance to nearest: quarry (a,c), canal (b), and road (e)
Hydrogeologic	Water table elevation (b,c,d), water table gradient (b,d), electrical conductivity (b), sulphate content (b,d), total dissolved solids (b), gysum saturation (b), aquifer fluctuations (e)
Geomorphologic	Elevation (b), slope (b), aspect (b), geomorphologic unit (b), distance to paleochannels (b), terrace levels (d)

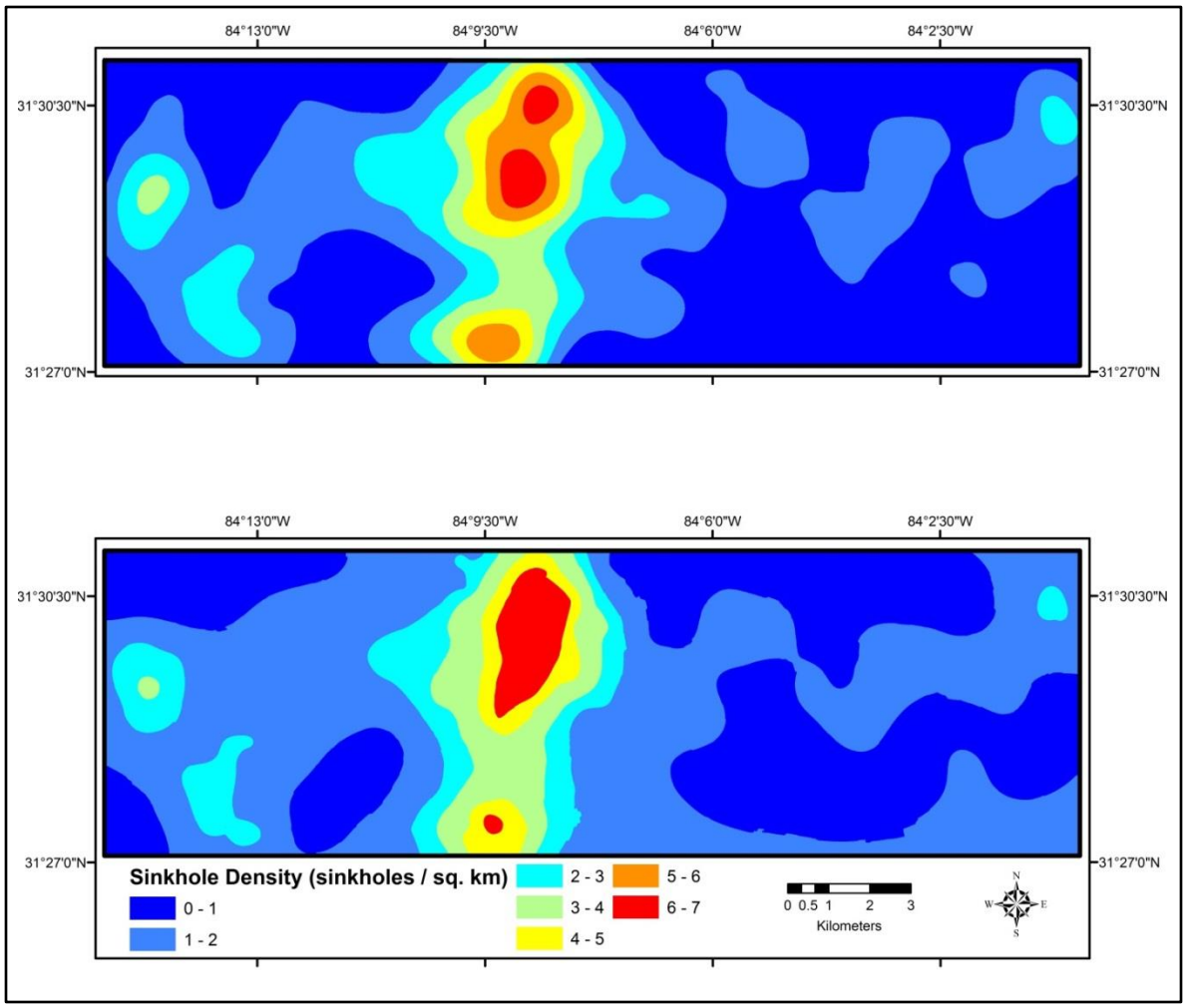

Figure 4.1: Measured (top) versus predicted (bottom) TD sinkhole density measurements. GWR prediction results were interpolated using the kriging method (RMS = 0.39).

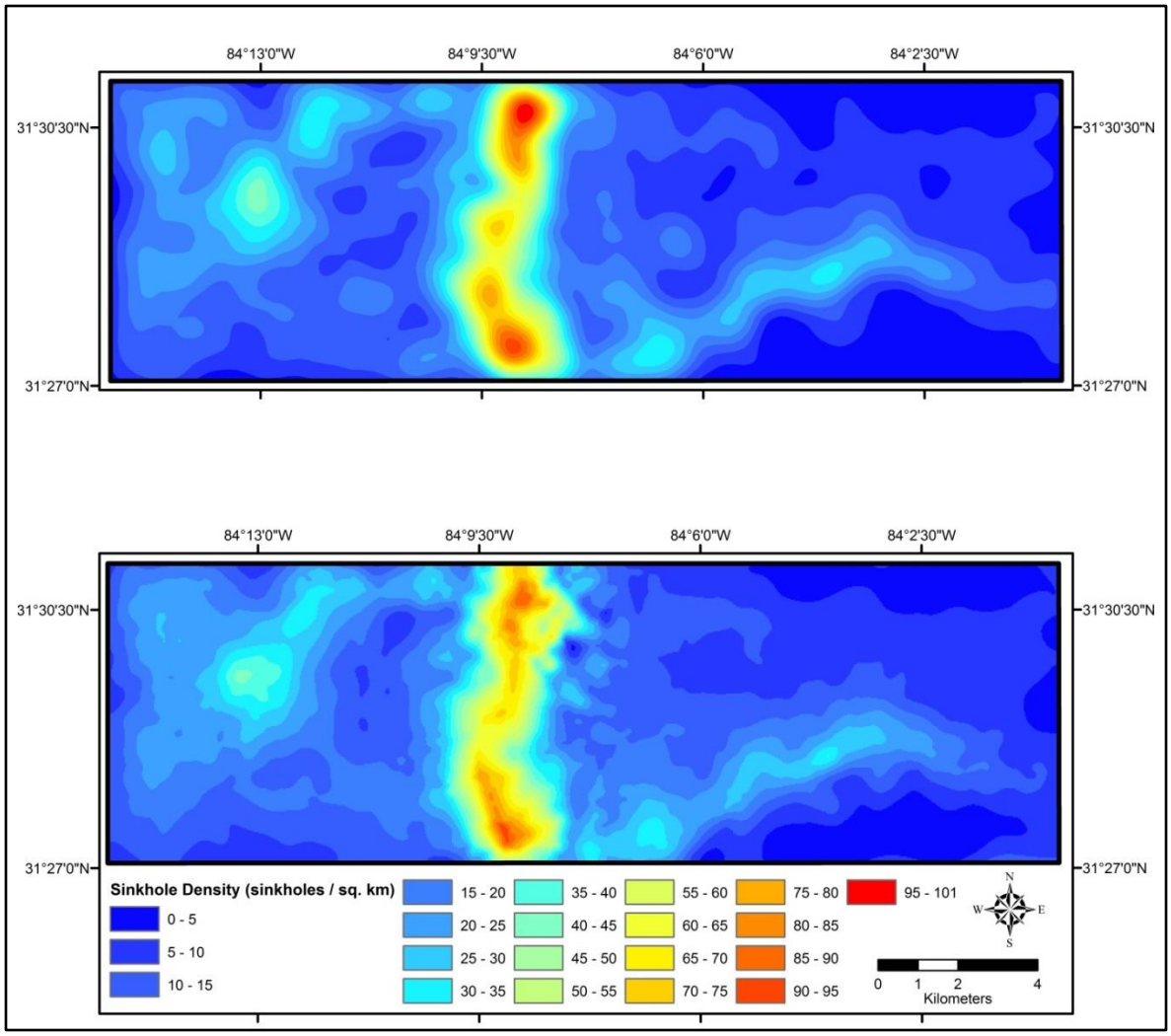

Figure 4.2: Measured (top) versus predicted (bottom) LiDAR sinkhole density measurements. GWR prediction results were interpolated using the kriging method (RMS = 1.79).

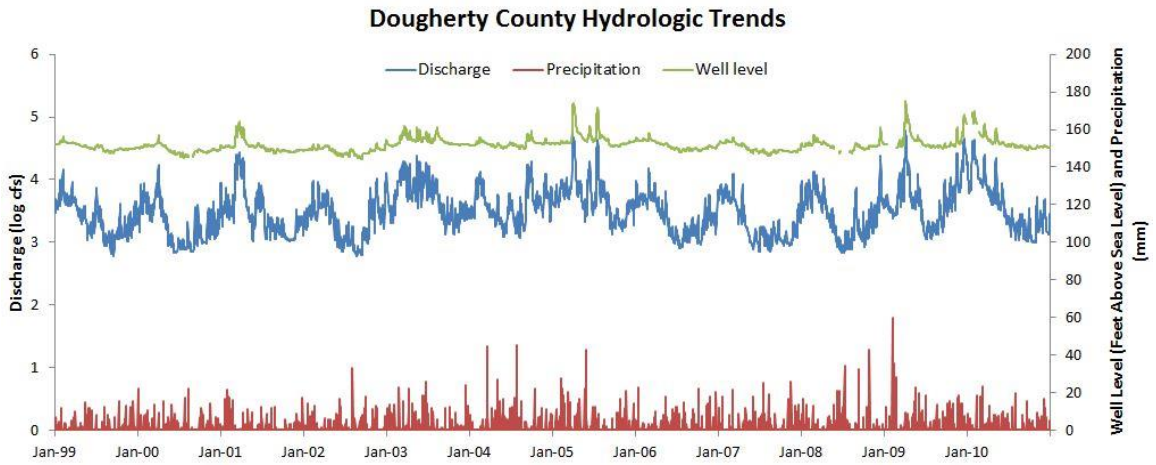

Figure 4.3: Time-series analysis of Flint River discharge, precipitation, and groundwater levels from 1999 – 2011.

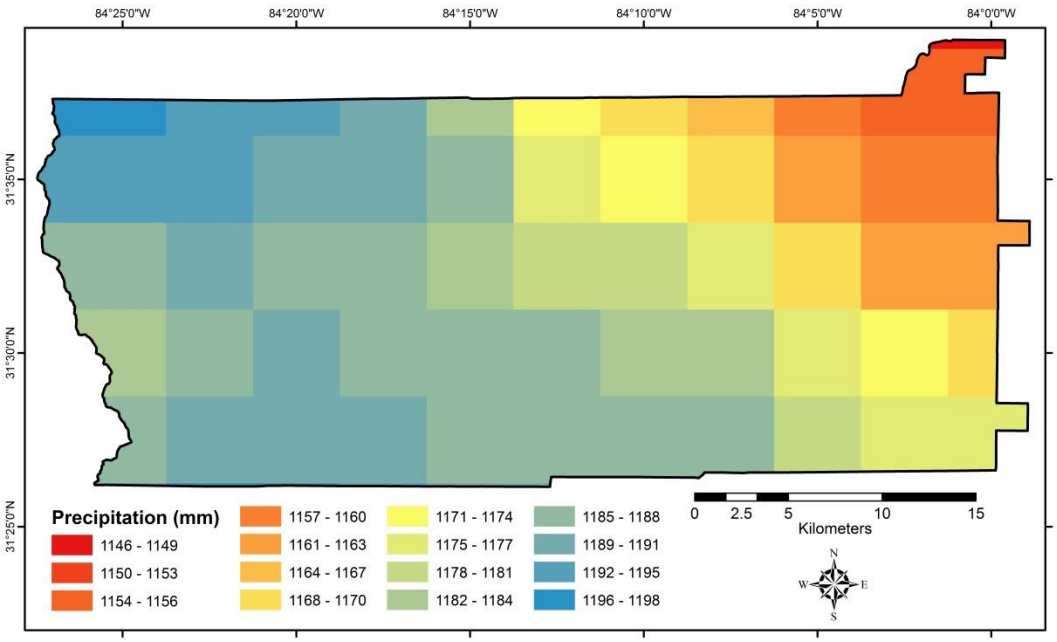

Figure 4.4: PRISM averaged annual precipitation for the period 1999 – 2011. The study area is outlined in southeast Dougherty County.

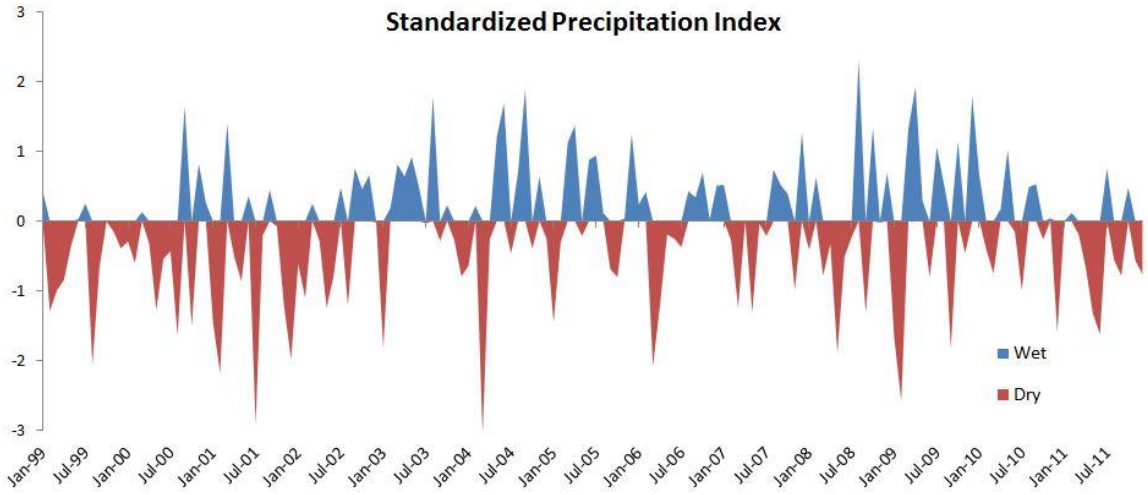

Figure 4.5: 1-month Standardized Precipitation Index derived from a National Oceanic and Atmospheric Administration (NOAA) gage located in central Dougherty County, Georgia.

CHAPTER 5

CONCLUSIONS

This paper integrated a GIS-based sinkhole mapping methodology with geostatistical techniques to analyze the primary factors and mechanisms influencing sinkhole formation within the mantled karst terrain of Dougherty County, Georgia. The results of our sinkhole formation models, including geologic, hydrologic, geomorphologic, hydrogeologic, and anthropogenic factors, show that hydrogeologic, hydrologic, and geologic variables are the most influential factors controlling sinkhole formation in both sinkhole datasets (i.e., TD and LiDAR). Specifically, areas with the following characteristics have a higher sinkhole formation probability within the study area: high aquifer fluctuations (12 – 14m), shallow overburden thickness (<7.5m), and close proximity to streams, wetlands, and bedrock fractures. These results were used to produce prediction maps for both sinkhole datasets.

The investigations of the present paper resemble aspects of previous work (Doctor and Doctor, 2012; Doctor *et al.*,2008b), in which geographically weighted regression was used to evaluate factors influencing sinkhole formation. This contribution examines a larger study area with multiple dependent variable datasets and independent variables that encompass many categories of sinkhole formation factors.

This study demonstrated that the application of spatial statistical techniques to multiple sinkhole inventory datasets (i.e., spatiotemporal and high-resolution) is an

effective approach to understanding sinkhole development. Applications of the demonstrated methodologies may be used by land use planners, water resource managers, and infrastructure developers to provide first-order estimates of future sinkhole occurrence to mitigate risks associated with sinkhole formation.

CHAPTER 6

DOUGHERTY COUNTY ASSESSMENT

The application of the methodologies employed in the southern Dougherty County study area was evaluated by implementing the sinkhole mapping and geostatistical-based sinkhole formation model across the entire extent of Dougherty County. This additional study provided a broader understanding of sinkhole formation in Dougherty County.

Beyond the different scale in the subsequent analysis, different data were utilized for the Dougherty County assessment. A Category-1 proposal approved by the European Space Agency (ESA) permitted European Remote Sensing (ERS) 1 and 2 satellite aperture radar (SAR) data to be utilized for a time-series of elevation data. Five high-resolution (5m) DEMs were generated from SAR data using the DEM extraction tool in ENVI's SARscape program. The 5m DEM's were from 1999, 2002, 2004, 2005, and 2009. Although they were analyzed separately from the SAR-generated DEMs due to spatial resolution differences, 30m DEM data from NASA's Terra satellite (ASTER sensor) and the SRTM were utilized to measure sinkhole development. These seven DEM's were acquired to analyze the spatiotemporal distribution of cover-subsidence (large-scale) and cover-collapse (small-scale) sinkhole development. The ancillary datasets utilized in the geostatistical-based sinkhole formation models were the same as the local study area, with the addition of the PRISM annual precipitation and a land cover

change product from the 2006 National Land Cover Database (NLCD). PRISM data could be included in the county-wide assessment due to a higher amount of variance over the larger study area.

Sinkhole inventory maps were created from each DEM. Temporal-difference (TD) point maps and subsequent density maps were generated by comparing the five 5m DEMs for cover-collapse sinkhole formation analysis and two 30m DEMs for cover-subsidence sinkhole formation analysis. The TD maps represented newly formed or enlarged sinkholes throughout the study period (e.g., 1999 – 2011). In total, 2,913 TD cover-collapse and 579 TD cover-subsidence sinkholes were identified in Dougherty County. In the local, southern Dougherty County study, 275 sinkholes were identified in the TD dataset. The difference in the number of mapped sinkholes between the local study and county-wide study is a result of three aspects: 1) the study area size, 2) the spatial resolution differences in the DEMs used to map sinkholes, and 3) the temporal resolution of the DEM data. 5m DEM data from 1999, 2002, 2004, 2005, and 2009 and 30m DEM data from 2000 and 2011 were used to produce the county-wide TD sinkhole dataset, while 10m data from only 1999 and 2011 were used to generate the local study area TD sinkhole dataset. This illustrates that the sinkhole maps should be treated as minimum sinkhole inventories due to spatial and temporal data resolution constraints.

Overall, the county-wide sinkhole formation models did not perform as well as the local models. Respectively, 79% and 64% of cover-collapse and cover-subsidence sinkhole distributions were explained by the sinkhole formation models, as measured by the GWR adjusted R^2 values. Similar to the local model, the addition of geochemical, pedologic, and other variables would improve the explanatory power of the sinkhole

formation model. When compared to the local sinkhole formation model, the lower adjusted R^2 values of the county-wide sinkhole formation model may reflect different sinkhole formation processes, mechanisms, and environmental conditions operating in the areas of Dougherty County not covered by the local model (e.g., western and northeastern Dougherty County). For example, clayey soils are more predominant and urban influences are minimal in western Dougherty County. Additionally, geostatistical methods are largely affected by the size of the study area, and it is more difficult to fit a model to a larger study area.

The GWR-prediction results for sinkhole density were input into deterministic interpolators and compared to the measured sinkhole density to test the ability of GWR to predict sinkhole distribution (Figures 6.1 and 6.2). The comparisons between the two sinkhole datasets yielded similar conclusions as the local study area sinkhole datasets. Though not as close as the local predictions, the county-wide results compare well as a first-order estimate. However, they tend to over-estimate sinkhole density, particularly in areas of high-sinkhole density and are not able to capture local heterogeneity.

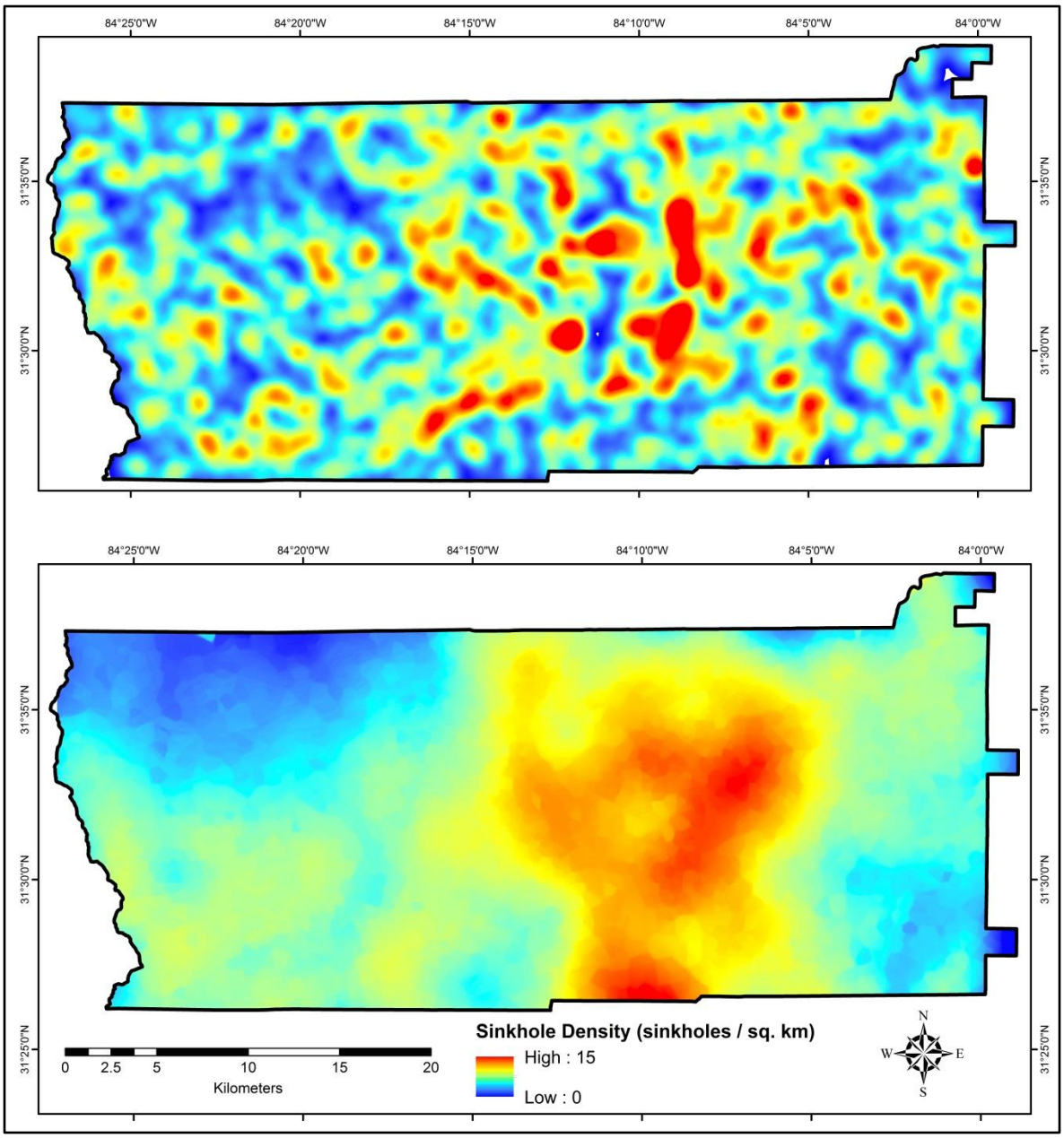

Figure 6.1: Measured (top) versus predicted (bottom) cover-collapse sinkhole density measurements. GWR prediction results were interpolated using a combination of inverse distance weighting and kriging methods (average RMS = 0.27).

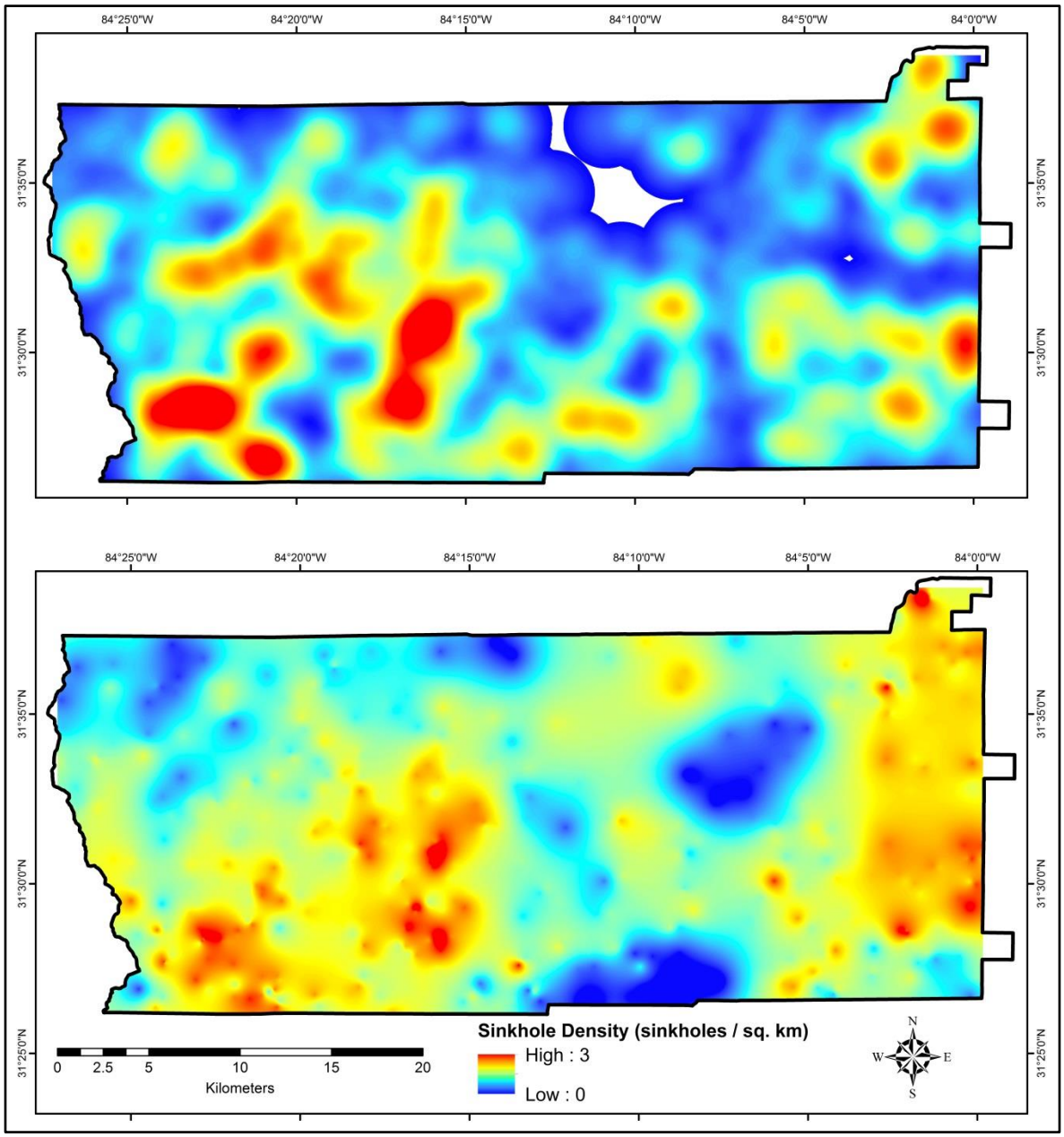

Figure 6.2: Measured (top) versus predicted (bottom) cover-subsidence sinkhole density measurements. GWR prediction results were interpolated using a combination of inverse distance weighting and kriging methods (average RMS = 0.36).

CHAPTER 7

SUMMARY

This thesis demonstrates the effectiveness of combining a GIS-based sinkhole mapping procedure with geostatistical techniques to understand spatiotemporal sinkhole development at varying scales in a mantled karst setting. The results provide insight into the specific factors and mechanisms most influential in sinkhole formation in Dougherty County, Georgia. The geostatistical prediction results have been shown to provide first-order estimates of sinkhole distribution based on the comparisons between the measured and predicted sinkhole density maps. Although local heterogeneity was not fully captured in all areas by the geostatistical prediction results, areas of high and low sinkhole density were successfully located.

The sinkhole formation analysis is presented by describing and comparing two components. In the first component, the sinkhole mapping, geostatistical, and prediction methodologies were developed in a local area within Dougherty County to quantify the effect of geologic, hydrologic, anthropogenic, hydrogeologic, and geomorphologic variables on sinkhole formation using spatiotemporal and high-resolution, primarily satellite-based datasets. The second component utilized the tested methodologies to understand and predict sinkhole development on a larger scale with different datasets.

The local and county-wide study were compared to further our understanding of sinkhole formation dynamics in Dougherty County. The variables with the greatest

explanatory power for county-wide sinkhole formation were similar to the southern Dougherty County sinkhole datasets (e.g., TD and LiDAR). Distance to wetlands, overburden thickness, and aquifer fluctuations were statistically significant in all sinkhole datasets. Distance to fractures and rivers were statistically significant in all but one sinkhole dataset (cover-subsidence). These results are consistent with previous sinkhole studies conducted in Dougherty County. Hyatt and Jacobs (1996) found that small-scale (cover-collapse) sinkholes primarily formed within the Flint River floodplain along bedrock fractures. This research agreed with Warner Gordon's (2012) findings that groundwater fluctuations in the Upper Floridan Aquifer are linked to sinkhole formation. These results also agree with observations from Tihansky (1999) and Hicks *et al.* (1987) regarding higher sinkhole occurrence in areas of lower overburden thickness. Brook and Allison (1986) described many large-scale sinkholes forming in wetland areas, specifically wetland areas west of the Flint River. Correspondingly, this research found the highest density of large-scale sinkholes to be in wetland areas west of the Flint River. Wetland areas in the eastern portion of the county also contained higher sinkhole densities.

The sinkhole mapping methodology described coupled with the recent increase in data availability (e.g., LiDAR) across the United States allows for the sinkhole mapping, geostatistical, and prediction methodologies to be applied to other karst regions that experience active sinkhole development in an effort to better understand karst hydrology and geology. The broader impacts of these results potentially improve land use, water resources, and infrastructure development decision-making.

The sinkhole mapping procedure, adapted from previously published work, performed well when compared to previously mapped sinkhole distributions in Dougherty County, especially when considering the vertical and horizontal accuracy errors in the DEMs. Nevertheless, the mapping methodology and geostatistical techniques could be applied in other karst regions with current, field-derived sinkhole inventories to test their accuracies. Additionally, other methods for predicting sinkhole distribution should be explored that might provide the ability to account for local heterogeneity.

REFERENCES

Al-Kouri O, Al-Fugara AK, Al-Rawashdeh S, Sadoun B, Pradhan B. 2013. Geospatial Modeling for Sinkholes Hazard Map Based on GIS & RS Data. *Journal of Geographic Information System* **05** : 584–592. DOI: 10.4236/jgis.2013.56055

Brook GA, Allison TL. 1986. Fracture Mapping and Ground Susceptibility Modeling. In *Covered Karst Terrain: the Example of Dougherty County, Georgia*. *Proceedings of Symposium on Land Subsidence, Venice, Italy.* IAHS Publication 151, International Association of Hydrological Sciences: Wallingford; 595-606.

Brook GA, Sun C. 1982. Predicting the specific capacities of wells penetrating the Ocala aquifer beneath the Dougherty Plain, Southwest Georgia. *Technical Completion Report USDI/OWRT Project A-086-GA* : 1–86.

Brunsdon C, Fotheringham S, Charlton M. 1998. Geographically Weighted Regression-modelling spatial non-stationarity. *Journal of the Royal Statistical Society: Series D (The Statistician)* 47 : 431–443. doi: 10.1111/1467-9884.00145

Clarke JS, Faye RE, Brooks R. 1984. Hydrogeology of the Clayton Aquifer of Southwest Georgia. *Georgia Department of Natural Resources, Georgia Environmental Protection Division, and Georgia Geologic Survey Accelerated Ground Water Program Hydrologic Atlas 13*.

Doctor DH, Doctor KZ. 2012. Spatial analysis of geologic and hydrologic features relating to sinkhole occurrence in Jefferson County, West Virginia. *Carbonates Evaporites* **27** : 143–152. DOI: 10.1007/s13146-012-0098-1

Doctor DH, Weary DJ, Orndorff RC, Harlow JGE, Kozar MD, Nelms DL. 2008a. Bedrock Structural Controls on the Occurrence of Sinkholes and Springs in the Northern Great Valley Karst, Virginia and West Virginia. *Sinkholes and the Engineering and Environmental Impacts of Karst* : 12-22. DOI: 10.1061/41003(327)2

Doctor DH, Young JA. 2013. An Evaluation of Automated GIS Tools for Delineating Karst Sinkholes and Closed Depressions from 1-Meter LiDAR-Derived Digital Elevation Data. In *Proceedings of the 13th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst* : 449-458.

Doctor KZ, Doctor DH, Kronenfeld B, Wong DWS, Brezinski DK. 2008b. Predicting Sinkhole Susceptibility in Frederick Valley, Maryland Using Geographically Weighted Regression. *Sinkholes and the Engineering and Environmental Impacts of Karst* : 243 – 256. DOI: 10.1061/41003(327)24

ESRI (Environmental Systems Resource Institute). 2012. ArcMap 10.1: ESRI, Redlands, California

Ford DC, Williams PW. 1989. Karst Geomorphology and Hydrology. Chapman & Hall: London

Ford DC, Palmer AN, White WB. 1988. Landform development; karst. In *Hydrogeology: The Geology of North America*, Back W, Rosenshein JS, Seaber PR (eds). Geological Society of America: Boulder, Colorado; 401-412.

Galve JP, Bonachea J, Remondo J, Gutiérrez F, Guerrero J, Lucha P, Cendrero A, Gutiérrez M, Sánchez J. 2008. Development and validation of sinkhole susceptibility models in mantled karst settings. A case study from the Ebro valley evaporite karst (NE Spain). *Engineering Geology* **99** : 185–197. DOI: 10.1016/j.enggeo.2007.11.011

Galve JP, Gutiérrez F, Lucha P, Guerrero J, Bonachea J, Remondo J, Cendrero A. 2009a. Probabilistic sinkhole modelling for hazard assessment. *Earth Surface Processes and Landforms* **34** : 437–452. DOI: 10.1002/esp.1753

Galve JP, Gutiérrez F, Remondo J, Bonachea J, Lucha P, Cendrero A. 2009b. Evaluating and comparing methods of sinkhole susceptibility mapping in the Ebro Valley evaporite karst (NE Spain). *Geomorphology* **111** (3-4) : 160–172. DOI: 10.1016/j.geomorph.2009.04.017

Gao Y, Alexander, Jr EC. 2007. Sinkhole hazard assessment in Minnesota using a decision tree model. *Environmental Geology* **54** : 945–956. DOI: 10.1007/s00254-007-0897-1

Gesch D, Oimoen M, Greenlee S, Nelson C, Steuck M, Tyler D. 2002. The National Elevation Dataset. *Photogrammetric Engineering and Remote Sensing* **68** (1) : 5-11.

Gutiérrez F, Galve JP, Guerrero J, Lucha P, Cendrero A, Remondo J, Bonachea J, Gutiérrez M, Sánchez JA. 2007. The origin, typology, spatial distribution and detrimental effects of the sinkholes developed in the alluvial evaporite karst of the Ebro River valley downstream of Zaragoza city (NE Spain). *Earth Surface Processes and Landforms* **32** : 912–928. doi: 10.1002/esp.1456

Gutiérrez F, Galve JP, Lucha P, Castañeda C, Bonachea J, Guerrero J. 2011. Integrating geomorphological mapping, trenching, InSAR and GPR for the identification and characterization of sinkholes: A review and application in the mantled evaporite karst of the Ebro Valley (NE Spain). *Geomorphology* **134** : 144–156. DOI: 10.1016/j.geomorph.2011.01.018

Gutiérrez M. 2013. Karst Geomorphology. In *Geomorphology*. CRC Press/Balkema: Leiden, The Netherlands; 179–227.

Hicks DW, Gill HE, Longsworth SA. 1987. Hydrogeology, chemical quality, and availability of ground water in the Upper Floridan Aquifer, Albany area, Georgia. *U.S. Geologic Survey Water Resources Investigations Report 87-4145* : 1–40.

Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G, Coulston J, Herold ND, Wickham JD, Megown K. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing* **81** : 345-354.

Hubbard DA. 2001. Sinkhole distribution of the Valley and Ridge Province, Virginia. In *Proceedings of the Eighth Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst* : 33-36.

Hyatt JA, Jacobs PM. 1996. Distribution and morphology of sinkholes triggered by flooding following Tropical Storm Alberto at Albany, Georgia, USA. *Geomorphology* **17** : 305–316. DOI: 10.1016/0169-555x(96)00014-1

Hyatt JA, Wilson R, Givens JS, Jacobs PM. 2001. Topographic, geologic, and hydrogeologic controls on dimensions and locations of sinkholes in thick covered karst, Lowndes County, Georgia. In *Proceedings of the Eighth Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst* : 37-45.

Kidner D, Dorey M, Smith D. 1999. What's the point? Interpolation and extrapolation with a regular grid DEM. In *Proceedings of the IV International Conference on GeoComputation*.

Lamelas MT, Marinoni O, Hoppe A, de la Riva J. 2008. Doline probability map using logistic regression and GIS technology in the central Ebro Basin (Spain). *Environmental Geology* **54** : 963–977. DOI: 10.1007/s00254-007-0895-3

MacNeil FS. 1947. Geologic map of the Tertiary and Quaternary Formations of Georgia. *U.S. Geological Survey Oil and Gas Inventory Map 72*.

McFadden SS, Perriello PD. 1983. Hydrogeology of the Clayton and Claiborne Aquifers in Southwestern Georgia. *Georgia Department of Natural Resources, Georgia Environmental Protection Division, and Georgia Geologic Survey Accelerated Ground Water Program Information Circular 55* : 1-59.

McSwain K. 1998. Hydrogeology of the Upper Floridan Aquifer in the Vicinity of the Marine Corps Logistics Base Near Albany, Georgia. *U.S. Geological Survey Water Resources Investigations Report 98-4202* : 1-49.

Newton JG. 1987. Development of Sinkholes Resulting from Man's Activities in the Eastern United States. *U.S. Geological Survey Circular 968* : 1-54.

Panno SV, Kelly WR, Angel JC, Luman DE. 2013. Hydrogeologic and topographic controls on evolution of karst features in Illinois' sinkhole plain. Carbonates Evaporites **28** : 13–21. DOI: 10.1007/s13146-013-0157-2

Parker EH, Hawman RB. 2012. Multi-channel Analysis of Surface Waves (MASW) in Karst Terrain, Southwest Georgia: Implications for Detecting Anomalous Features and Fracture Zones. *Journal of Environmental & Engineering Geophysics* **17** : 129–150. DOI: 10.2113/jeeg17.3.129

Rahimi M, Alexander, Jr. EC. 2013. Locating Sinkholes in LiDAR Coverage of a Glacio-Fluvial Karst, Winona County, MN. In *Proceedings of the 13th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst* : 469-480.

Rose MD, Federico A, Parise M. 2004. Sinkhole genesis and evolution in Apulia, and their interrelations with the anthropogenic environment. *Natural Hazards and Earth System Science* **4** : 747–755. DOI: 10.5194/nhess-4-747-2004

Shaw Faulkner MG, Stafford KW, Bryant AW. 2013. Delineation and Classification of Karst Depressions Using LiDAR: Fort Hood Military Installation, Texas. In *Proceedings of the 13th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst* : 459-468.

Stewart LM, Warner D, Dawson BJ. 1999. Hydrogeology and Water Quality of the Upper Floridan Aquifer, Western Albany Area, Georgia. *U.S. Geologic Survey Water Resources Investigations Report 99-4140* : 1–42.

Stokes T, Griffiths P, Ramsey C. 2010. Karst Geomorphology, Hydrology, and Management. In *Compendium of Forest Hydrology and Geomorphology in British Columbia*, Pike RG, Redding TE, Moore RD, Winkler, RD, and Bladon, KD (eds). Forest Science Program / FORREX; 373-400.

Tihansky, AB. 1999. Sinkholes, west-central Florida. In *Land Subsidence in the United States*, Galloway D, Jones DR, Ingebritsen SE (eds). U.S. Geological Survey Circular 1182 : 121-140.

Torak LJ, Davis GS, Strain GA, Herndon JG. 1993. Geohydrology and evaluation of water-resources potential of the Upper Floridan aquifer in the Albany area, southwestern Georgia. *U.S. Geological Survey Water Supply Paper 2391* : 1–59.

Torak LJ, Painter JA. 2006. Geohydrology of the Lower Apalachicola-Chattahoochee-Flint River Basin, Southwestern Georgia, Northwestern Florida, and Southeastern Alabama. *U.S. Geological Survey Scientific Investigations Report 2006-5070* : 1-73.

USFWS (United States Fish and Wildlife Service). 2002. National Wetlands Inventory website. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. http://www.fws.gov/wetlands/

Waltham T, Bell FG, Culshaw MG. 2005. *Sinkholes and Subsidence: Karst and Cavernous Rocks in Engineering and Construction*. Springer: Berlin

Warner D. 1997. Hydrogeologic Evaluation of the Upper Floridan Aquifre in the Southwestern Albany Area, Georgia. *U.S. Geological Survey Water Resources Investigations Report 97-4129* : 1-27.

Warner Gordon D, Painter JA, McCranie JM. 2012. Hydrologic Conditions, Groundwater Quality, and Analysis of Sinkhole Formation in the Albany Area of Dougherty County, Georgia, 2009. *U.S. Geological Survey Scientific Investigations Report 2012-5018* : 1-60.

White EL, Aron G, White WB. 1986. The influence of urbanization of sinkhole development in central Pennsylvania. *Environmental Geology and Water Sciences* **8** : 91–97. DOI: 10.1007/bf02525562

Wilson WL, Beck BF. 1992. Hydrogeologic Factors Affecting New Sinkhole Development in the Orlando Area, Florida. *Groundwater* **30** : 918–930. DOI: 10.1111/j.1745-6584.1992.tb01575.x

APPENDIX A

SINKHOLE FORMATION FACTOR PRODUCTION

This appendix details the production of the interpolated independent variables used in the sinkhole formation geostatistical models. The Geostatistical Analyst toolbar within ESRI's ArcGIS 10.1 was used for interpolating the point data.

Upper Floridan Aquifer Fluctuations

Daily data for 12 USGS observation wells installed in the Upper Floridan Aquifer and actively operated during the study period were downloaded. The difference between the maximum and minimum well level during each year between 1999 and 2011 was computed. Then the average of those values was found and assigned to the georeferenced point used for interpolation. Figure A.1 shows locations of the wells used to produce the interpolated surface. Figure A.2 displays the relationship between Upper Floridan Aquifer fluctuations and proximity to bedrock fractures. It is known that specific yields of groundwater wells increase greatly within approximately 30m of a fracture in the study area. However, these data do not show a strong correlation between fluctuation level and proximity to fractures like specific yield and proximity to fractures do.

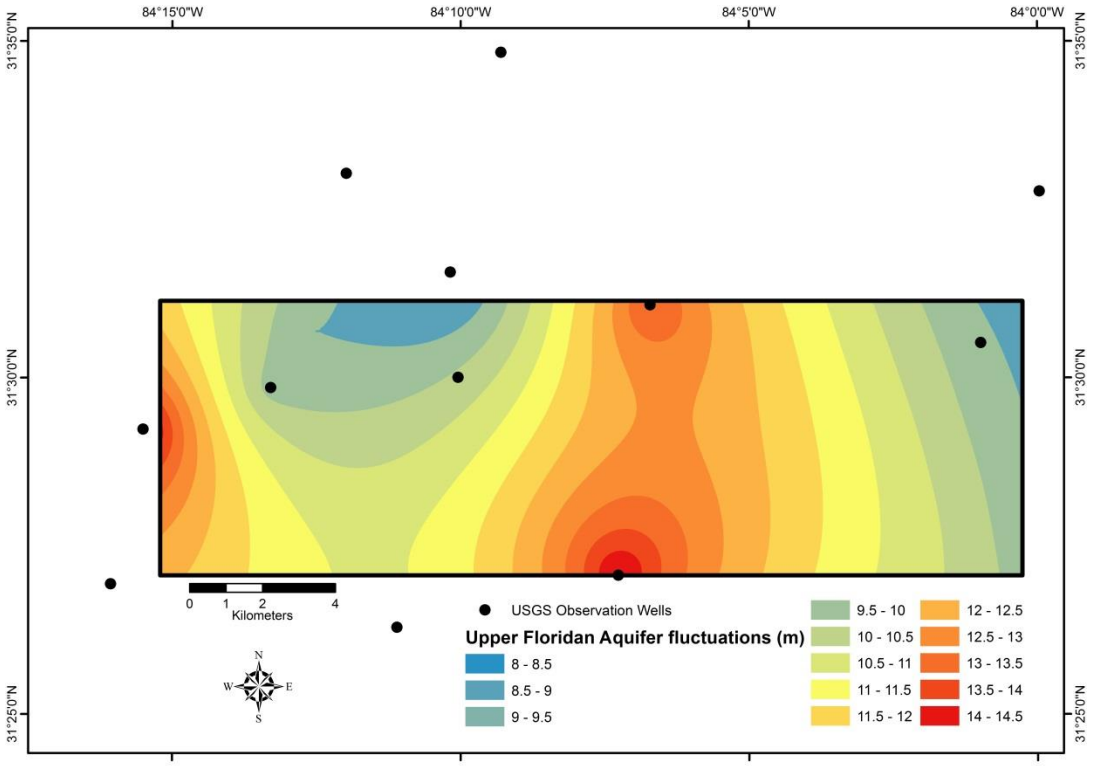

Figure A.1: Upper Floridan Aquifer fluctuation map calculated using groundwater level data from USGS wells covering the entire study period.

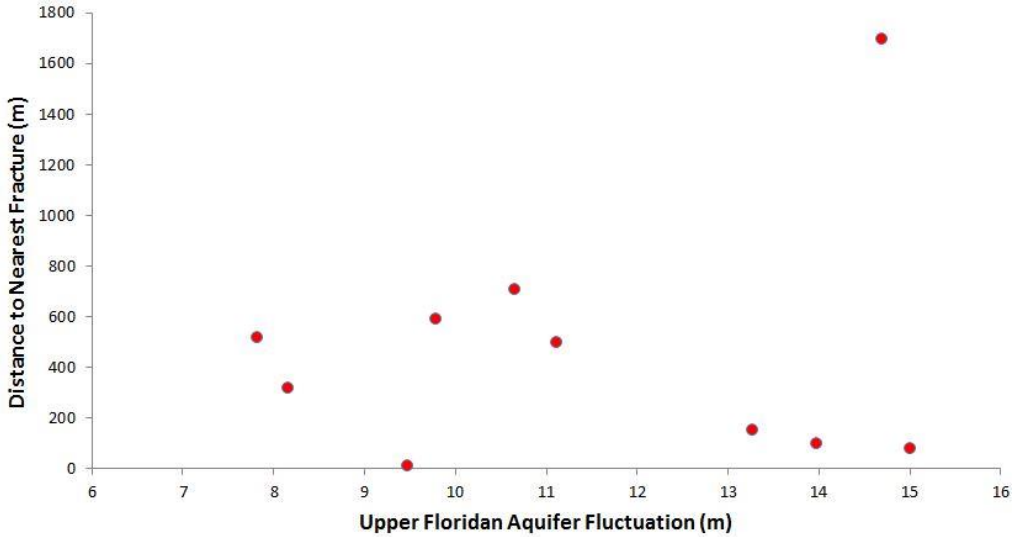

Figure A.2: Scatterplot showing the relationship between Upper Floridan Aquifer fluctuations and proximity to bedrock fractures.

Overburden Thickness

Overburden thickness (i.e., depth to bedrock) measurements were gathered from geophysical data, borehole analysis, and subsurface cross sections for 33 locations. 15 more locations were added along the Flint River based on knowledge of the Ocala Limestone exposure along the Flint River due to erosional processes. The georeferenced point data were input into an Empirical Bayesian Kriging interpolator to produce a smooth surface clipped to the study area (RMS = 2.42).

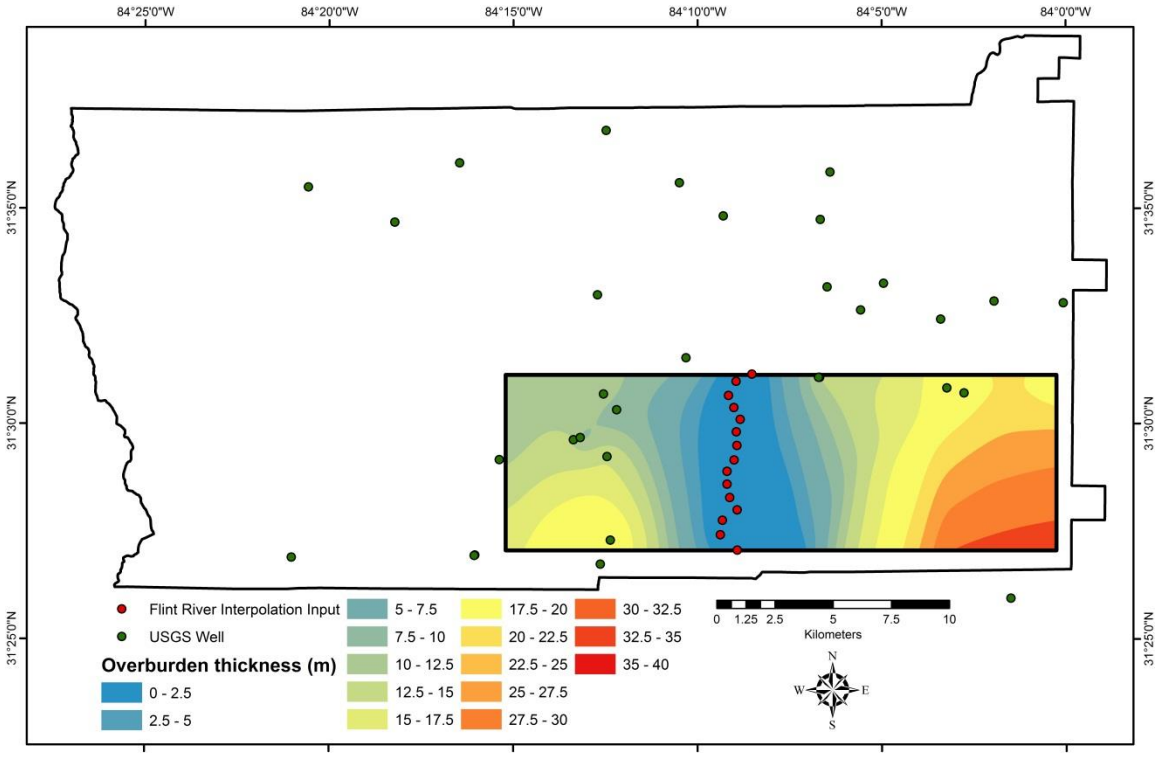

Figure A.3: Overburden thickness (i.e., depth to bedrock) map calculated using borehole data from USGS studies.