**A METHOD OF MAPPING SINKHOLE SUSCEPTIBILITY
USING A GEOGRAPHIC INFORMATION SYSTEM:
A CASE STUDY FOR INTERSTATES IN
THE KARST COUNTIES OF VIRGINIA**

By

Alexandra L. Todd

Lindsay Ivey-Burden, Advisor
Department of Civil Engineering

Thesis submitted to the faculty of the University of Virginia
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
In
Civil Engineering

Thornton Hall B228
351 McCormick Road,
P.O. Box 400742
Charlottesville, VA, 22904
(504)982-4620
alt5ry@virginia.edu

November 19, 2014

A METHOD OF MAPPING SINKHOLE SUSCEPTIBILITY
USING A GEOGRAPHIC INFORMATION SYSTEM:
A CASE STUDY FOR INTERSTATES IN
THE KARST COUNTIES OF VIRGINIA

Alexandra L. Todd

ABSTRACT

Karst terrain is landscape underlain by limestone that has been chemically dissolved by acidic groundwater, producing subsurface voids that pose risks for sinkholes if the overlaying soil can no longer support its own weight and collapses. The western counties of Virginia are heavy in karst due to their natural, geographic boundary of the western Ridge Province and the eastern Blue Ridge Mountain Range. As a result, the Commonwealth of Virginia Hazard Mitigation Plan recommends that the Virginia Department of Transportation (VDOT) develop a method to determine the roadways and regions most susceptible to experiencing sinkholes, in an effort to reduce the number of reported sinkhole damage to property. While many noninvasive methods exist to detect subsurface voids, such as electric resistivity imaging, ground penetrating radar, and seismic surveys, these methods are time consuming and costly.

This study proposes the use of a geographic information system (GIS) to create a susceptibility map, pinpointing regions in the karst counties of Virginia, in particular, along interstates, most susceptible to future sinkhole development, determined by five factors that have previously been shown to play a role in the acceleration of sinkhole formation in Virginia: bedrock type, proximity to fault lines, drainage class, slope of incised river banks, and minimum soil depth to bedrock. The analysis compares a 1:24,000 scale map of existing sinkholes developed by Virginia Department of Mines Minerals and Energy (DMME) geologist, David Hubbard, with a series of risk maps representing differing combinations of each of the five risk factors to determine which weighted combination is most appropriate to use for a final representative risk map. The layers representing each risk factor are created using publicly available tabular and spatial data taken from the USDA Soil Survey Geographic (SSURGO) Database, the USGS National Map, the USGS Mineral Resources Online Data, and the National Weather Service. The final combination choice will provide an idea of the corresponding factor's influence on predicting sinkhole risk regions. This investigation identified the following results for karst terrain in Virginia: (1) bedrock type has the most significant impact on predicting sinkhole risk, (2) proximity to faults plays a minimal, yet present, role in determining sinkhole risk, (3) drainage class is the second most influential factor in sinkhole formation behind bedrock type, (4) slope of incised river bank plays no role in the formation of sinkholes in Virginia, and (5) depth of overlying soil to bedrock has an existent yet insignificant effect on sinkhole development. The results display how this new inexpensive and efficient method of predicting sinkhole susceptibility can highlight the influence of natural features that trigger sinkhole and provide a map that can be used by local transportation departments as a general guideline to visualize regions along heavily trafficked interstates most and least at risk for sinkhole collapse. A benefit to this methodology is that the new technique can be adjusted to accommodate for sinkhole susceptibility in regions across the world, by simply adjusting the input risk layers to consider sinkhole risk potential based on the specific geology of a particular region.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS:

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION
Karst terrain forms as groundwater interacts with soluble bedrock, during which subsurface drainage causes unique solutional patterns to carve into the rocks, forming cavities and triggering land subsidence in the event that topsoil subsides into the voids *(1)*. These funnel-shaped closed-contour depressions on the earth's surface *(2)*, known as sinkholes, form in the presence of bedrock rich in carbon ions and are highly vulnerable to acidic groundwater *(3)*. As acidic rain hits the earth's surface, the carbon dioxide within it percolates through the bedrock, disappearing into the ground and naturally dissolving and enlarging the fractures and fissures that are characteristic of carbonate rocks, through a process known as dissolution. As the underground openings in the carbonate rock grow, water transport increases, and the dissolution process accelerates *(4)*. Eventually, a void reaches the point where the soil overlaying it softens, ravels, and erodes away until it is no longer capable of supporting the weight of the heavy, wet soil above *(5)*. When this happens, the roof collapses and creates a sinkhole, which serves as the focus of this paper.

The western counties of the state of Virginia are heavy in karst, because they are bounded on the west by the Ridge Province in West Virginia and on the east by the Blue Ridge Mountain Range, ultimately locating them in a long valley containing extensive folds and fractures of limestone and dolomite bedrock *(6)*. This folded and faulted regional topography is the result of differential weathering of rock belts over geologic time, and provides a natural likelihood for karst terrain and sinkhole formation *(7)*.

If a region overlays carbonate rocks and has a greater depth to bedrock than depth to water table, significant fluctuations of the water table present a potential for subsidence. As the water table drops, overburden materials lose saturation and ultimately lose buoyancy, which may trigger collapse, or, if the water table rises, the rewetting of dry materials could reduce their shear strength and result in structure failure as well *(2)*.

Sinkholes pose the risk of damaging property and endangering lives if developed in a highly populated or well-traveled area. While in a stable, free-of-development site, sinkholes are considered natural events, it is important to note that, often, human interaction tweaks and triggers the variation in water table levels more than the natural environment would. As sites continue to be developed, subsurface water systems are further perturbed, in an unintentional yet costly act *(8)* through the process of groundwater pumping, modification of natural drainage patterns, and inappropriate storm water management affecting local groundwater tables and consequently accelerate and spatially concentrate sinkhole subsidence (*6*). However, anthropogenic factors to subsidence hazards are outside the scope of this paper, and we will remain focused on naturally prone sinkhole regions and how we can pinpoint at-risk regions from this point forward.

This paper focuses on the *natural* factors of sinkhole formation and how, while impossible to fully eliminate natural karst hazards, losses and damages can be alleviated through effectively implementing certain techniques (*9*). The investigation proposes a new method of using a geographic information system (GIS) to predict where those sinkholes might form in an effort to avoid such dangers specifically along Virginia highways. Because of the public availability of spatial and tabular datasets provided by agencies such as the United States Geological Survey (USGS), the use of GIS techniques has become significantly useful to state and local governments in the field of natural hazards (*10*).

MOTIVATION
Roadways overlying regions of karst terrain are at risk of structural damage due to sinkhole formation underneath. Sinkholes have proven in the past to be destructive to life and property, leading to construction-site messes, contamination of groundwater, troublesome shifts and misalignments in infrastructure, and, of biggest concern to the Virginia Department of Transportation (VDOT), collapse and subsidence of the earth's surface near roadways (*8*)*.* In 2011, in Rockbridge County, Virginia, northbound lanes of Interstate 81, located at mile marker 170, were closed because of a sinkhole that opened up directly under the roadway. In 2004, Tropical Storm Gaston led to the formation of a 30 foot sinkhole that swallowed a busy Richmond, Virginia intersection. In 1910, in the city of Staunton, three large sinkholes opened up, swallowing an entire house and a 35-foot tree, and killing one worker who fell into a fissure that opened up as a result of the largest sinkhole (*11*). These are just three of the many sinkholes that have opened up across Virginia, damaging property, threatening lives, and causing large repair costs for the state.

An estimated three to five of VDOT's districts come in contact with karst terrain during construction (*8*)*.* Solid limestone with a high bearing capacity contains inner voids, whose variability, especially at shallow depths, threatens the integrity of the foundation and the ability to excavate in such a region *(12)*. The Commonwealth of Virginia Hazard Mitigation Plan recommends that VDOT determine the roadways and regions most susceptible to experiencing sinkholes to help lessen the number of reported sinkhole damage to property (*11*). While a number of noninvasive methods exist to detect subsurface karst, such as electric resistivity imaging, ground penetrating radar (GPR), and seismic surveys, these methods can be expensive, time consuming, and ineffective *(8)*. This study proposes the use of a GIS to pinpoint those regions along Virginia interstates most at risk of future sinkhole development, determined by a number of factors that have previously been shown to play a role in the acceleration of sinkhole formation. The results provide an inexpensive and quick method of better locating proposed roadway passages in an effort to avoid impact to karst areas *(13)* and determining which roadways may require immediate safety evaluations, ultimately minimizing environmental threats to life and property and maximizing land use *(14)*. Road and infrastructure construction can be detoured around regions with high risk of subsidence, or built with the most careful consideration to local water table adjustments, to help reduce new sinkholes from forming and to be certain that, in the case of a sinkhole surfacing, no persons or property will be damaged.

BACKGROUND/PREVIOUS STUDIES
Before recent studies, sinkhole risk could not be accurately estimated because of the lack of detailed mapping and historical records on the subject. Karst maps were the sole method of assessing subsidence probability. However, the Commonwealth of Virginia Hazard Mitigation Plan claims that "a high percentage of karst geology in a jurisdiction does not necessarily [imply] that the whole locality is at risk for land subsidence" (*11*). Without a well-established set of guidelines that predict probabilities, a true risk determination cannot be formed.

In a 2001 study on sinkhole distribution in Virginia, David A Hubbard, Jr. determined sinkhole locations by stereoscopic viewing and panchromatic aerial photography, field-checking any questionable sinkholes. Hubbard was able to determine that "sinkhole distribution across Virginia's Valley and Ridge province is not random," observing "three obvious factors in the distribution of karst as expressed by sinkholes: lithology, structure, and hydraulic gradient" *(1)*. Detected sinkholes mainly occur in regions where carbonate rocks are present, where structural

folds and faults exist, and adjacent to deep rivers and tributaries. However, Hubbard noted that not all sinkholes can be detected by remote sensing methods, since aircraft tilt makes certain shallow sinkholes entirely unrecognizable while making other low slope regions appear as sinkholes when in fact they were not. In fact, he estimated that it would take 250 years to map every single sinkhole in Virginia's Valley and Ridge province using solely aerial photography and field-checks (*1*).

This claim drove many geotechnical engineers to look further into the subject to develop trends that could be implemented into a GIS to create more accurate and efficient risk maps. Hyland statistically tested Hubbard's hypothesis using ArcMap, concluding that sinkholes have proven to form in regions of relatively pure carbonate rocks, such as limestone and dolomite, of the Ordovician age. She verifies that there is a correlation of sinkhole proximity with existing fault lines, which explains why the most significant karst terrains in Virginia are located in the Shenandoah Valley, a region containing extensive folds and fractures of limestone and dolomite bedrock (*4*).

Hyland also determined that sinkholes occur most sparsely near streams, and are most abundant approximately 600 to 1400 feet away from surface streams (*4*). While Hyland's results regarding proximity to surface streams might be statistically proven for *her* region of interest, it is risky to make generalizations about results of a karst technology from area to another because of the differences in lithology even between neighboring regions. This has become a precarious issue in GIS analyses as it becomes easy to get caught up in repeating the same tricks and results used for one location in a second location, when often correlating results does not necessarily apply to those other locations.

In determining regions heavy with karst, drainage says a lot about how at risk a site may be for sinkhole formation. Water flow is not manifested on the surface as much as in a non-karst region because it percolates into the subsurface caves. Thus smaller surface streams often do not exist or endure since subsurface drainage will swallow most of the water into the fissures, leaving only the stronger, more heavily flowing rivers to remain above ground. When it comes to surface streams and sinkhole risk, it makes more sense to say that proximity to deeply incised rivers plays a larger role in sinkhole formation than a specific distance from just *any* surface stream. Hubbard defended this idea by attributing the higher risk near incised streams to the steepened hydraulic gradient and the resulting increase in groundwater flow for those areas (*15*). If water flows through a deeply carved out channel, then there is essentially a very high relief zone adjacent to the channel, allowing a larger surface area for significant amounts of water to percolate into the neighboring bedrock.

Since a majority of the 26 most significant karst-containing counties found along the western edge of Virginia are *entirely* underlain by karsted carbonate bedrock, Green, Marken, et al. decided that sinkhole risk studies should focus on regions of bedrock within 75 feet of the surface, concluding that the timescale for which sinkholes may develop can be hours to months for shallow depth to bedrock, where it may be decades to centuries with a thicker depth to bedrock (*16*). Doctor et al. recommends a knowledge of both geologic structural information in a region and hydrologic data for a most thorough understanding of sinkhole prediction, after showing that the karst in the Great Valley formed largely due to aggressive and rising subsurface fluid (*17*). Doctor et al. came to the conclusion that clustering of sinkholes served as a proven sinkhole predictor (*18*). Similarly, a study by Zhou et al. identified a cluster of sinkholes along Interstate 71 in Maryland, and concluded that there is, in fact, an interaction between existing and new sinkholes, and the new sinkholes will most likely form within a 30 meter radius of

existing ones (*19*).

While Doctor and Zhou's claims of sinkholes forming in regions with existing clusters might be true, it is important to understand that this conclusion is only valid in a region with a stable environment. Construction in an area that currently has no known sinkholes can trigger new sinkhole formations, so we should only trust this prior claim if the cluster has occurred in a region with unchanging surrounding infrastructure. If this is the case, there must in fact be some subsurface aquifer or epikarstic drain nearby, leading to the natural cluster of sinkholes. Hence a region consisting of high paced drainage might imply a subsurface drainage source and could potentially tell us something about the possibility of sinkhole formation in that particular area.

With these initial trends in mind, this study aims to create a sinkhole risk analysis map by combining 5 factors – bedrock type, proximity to fault lines, drainage class, proximity to incised river banks, and depth of the overlying soil– into a single representative map spanning the western counties of Virginia.

METHODOLOGY
Study Area
Twenty-six Virginian counties contain karst terrain, composing a region that is bounded by Blue Ridge province on the east and by the West Virginian Appalachian Plateaus province on the west (*1*). The Virginia Department of Mines, Minerals, and Energy (DMME) includes Loudon county as a twenty-seventh because of sizeable amounts of limestone found in the sand and siltstone conglomerate, resulting in rock containing up to 50% carbonate and reported subsidence issues and karst flooding. The region of interest in this study involves these 27 counties (***Figure 1***).

The region includes approximately 29,853 square kilometers and ranges from -83°40'32'' to –77°19'42'' latitudinally and from 39°27'57'' to 36°35'37'' longitudinally in the Geographic Coordinate System, and the 1983 North American Datum.

Figure 1: Virginia Counties in Study Area
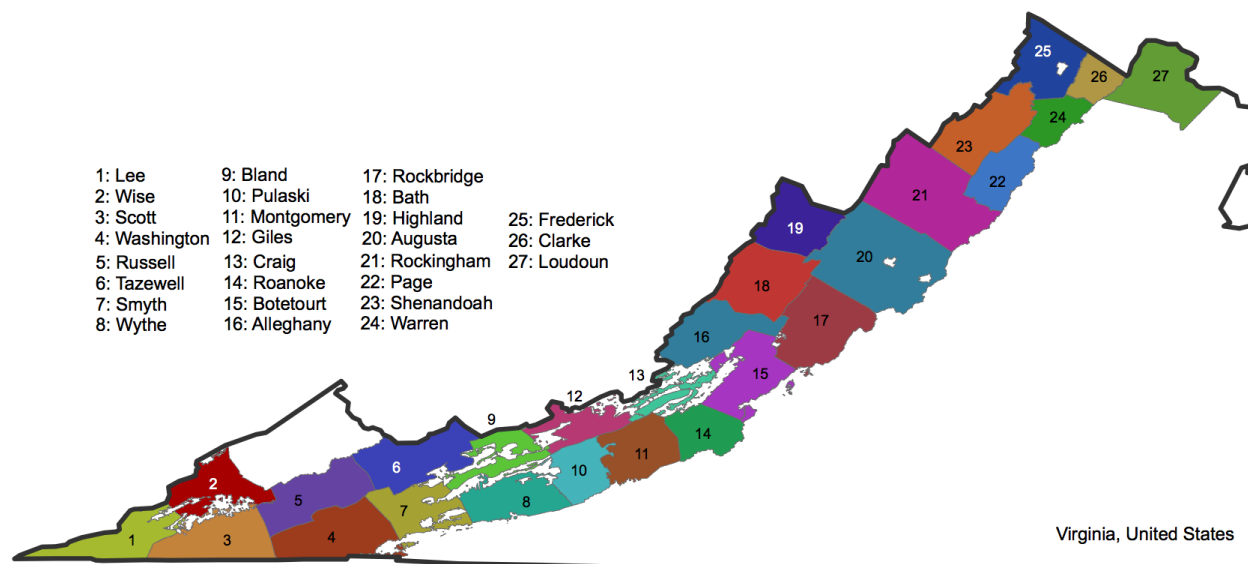
Data Acquisition and Preparation
The data used was taken from four different sources (***Table 1***). The analysis created a compiled

ranked map determining regions in danger of potential sinkhole formation based on five unique risk layers created in ArcMap 10.1. Each layer contained a map of regions assigned a 'Risk Value' ranging from 1-15, representing the level of potential hazard based on the corresponding risk factor. The distinct maps were ultimately combined using specific weights representing the corresponding factor's influence on predicting sinkhole risk regions, statistically deciding which weight combination was the most appropriate to use for the final representative risk map.

Table 1: Sources for Data Used

Data Name	Data Type	Source
Bedrock Type	Tabular Data	Soil Survey Geographic Database (SSURGO), USDA
County and Map Unit Boundaries	Shapefile	Soil Survey Geographic Database (SSURGO), USDA
Depth to Bedrock	Tabular Data	Soil Survey Geographic Database (SSURGO), USDA
Digital Elevation Model	Raster	National Map, USGS
Drainage Classes	Tabular Data	Soil Survey Geographic Database (SSURGO), USDA
Fault Lines	Shapefile	Mineral Resources Online Spatial Data, USGS
Rivers	Shapefile	National Weather Service website
Virginia State Boundary	Shapefile	National Map, USGS

Bedrock Type
Bedrock type is a contributing factor to sinkhole formation since sinkholes have proven to form in regions of relatively pure carbonate rocks *(4)*. The bedrock type layer was derived from SSURGO tabular data located in the *Component,* the *Component Parent Material (COPM),* and the *Component Parent Material Group (COPMGRP)* tables for each of the 27 counties of interest. Desired attributes from the individual tables were combined into a single table based on common fields through a Python script (**Appendix: Figure 11**). The parent material kind and origin were defined in the *COPM* table, where formation kind was described as *"alluvium", "colluvium",* or *"residuum"*, and formation origin ranged through a number of bedrock origins, such as *limestone, sandstone,* or *shale*. The *Component* table defined the corresponding representative component percentage. The *COPMGRP* table included common fields that allowed the *Component* and *COPM* tables to join together first and then eventually with the map units for visual analysis.

The script removed any duplicate or null entries as well as any entries where the parent material kind was alluvial, since alluvium soils rarely play a role in karst development. If a sinkhole does in fact exist in a region with alluvial deposits, it is more likely a result of one of the other factors in this study (proximity to faults or depth to bedrock, for example) rather than a result of the bedrock type *(4)*.

A new field called the "Unweighted Sinkhole Value" was added to the resulting table, determined by the parent material origin, assigning pure limestone origins the highest value of 4, limestones and dolomite combinations a value of 3, values containing only partial carbonate rock a 2, and entirely clastic and non-carbonate origins the lowest value of 1. The final table from the script was imported into Microsoft Excel, where a final weighted Sinkhole Value was calculated per map unit using the component percentage and the Unweighted Sinkhole Value:

$$Component1\ Percentage\ x\ Unweighted\ Sinkhole\ Value1$$
$$+\ Component2\ Percentage\ x\ Unweighted\ Sinkhole\ Value2$$
$$+\cdots$$
$$=\ Final\ Sinkhole\ Value\ for\ Corresponding\ Map\ Unit$$

While the final weighted Sinkhole Values ranged from 1-4, the values were reclassified into 15 equally incremented categories to remain consistent with the next four layers. This final table was imported into a Microsoft Access Database to then be imported into ArcMap, where it could be joined spatially with the spatial map units using the "Add/Join" tool in ArcMap (**Figure 2**).

Figure 2: Bedrock Type Risk Layer Used in Weighted Combination
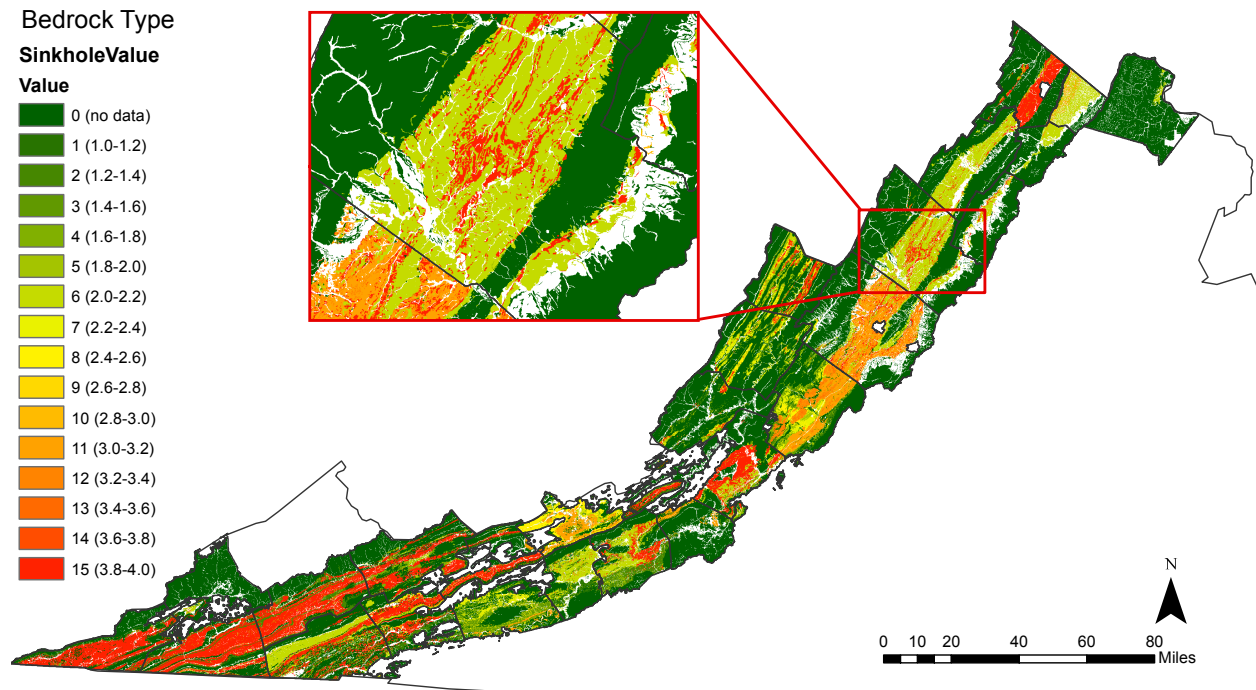
Proximity to Fault Lines

Previous studies have determined a correlation between sinkhole formation and proximity to existing fault lines, due to the fact that faults provide a quick route for ground water to percolate into. To create this fault risk layer in ArcMap, the "Multiple Ring Buffer" tool was used around the fault lines extracted from USGS Mineral Resources Online Spatial Data. Rings ranging from 0 – 3000 feet from the faults in increments of 200 feet were created and converted to raster, to then be reclassified into Sinkhole Values of 1 through 15. Values were highest at regions closest to the faults, decreasing outwardly as distance increased from faults (**Figure 3**).

Drainage Class

Drainage plays a role in predicting sinkhole risk because it provides information on how surface water behaves, whether manifesting on the surface in non-karst regions or draining into subsurface fissures in karstic regions. Drainage class (*drainagecl)* was a field defined in the SSURGO *Component* table, composed of values ranging from *excessively drained* to *very poorly drained*. A Python script was written creating a table that combined each map unit with its corresponding drainage class (**Appendix: Figure 12**). As done in previous layers, the table was imported into a Microsoft Access Database to then be added into ArcMap. Because there were only 7 different drainage classes listed, the regions with poor drainage had Sinkhole Values defined as the odd numbers ranging from 1 to 5, and regions that were excessively drained had

Sinkhole Values defined as the odd numbers ranging from 9 to 15, since this most likely meant water was being absorbed into subsurface karstic drainage systems. Zero was assigned to the excess regions with no drainage class assignment. Sinkhole Value 7 had no assignment in this layer. (**Figure 4**).

Figure 3: Proximity to Faults Risk Layer Used in Weighted Combination
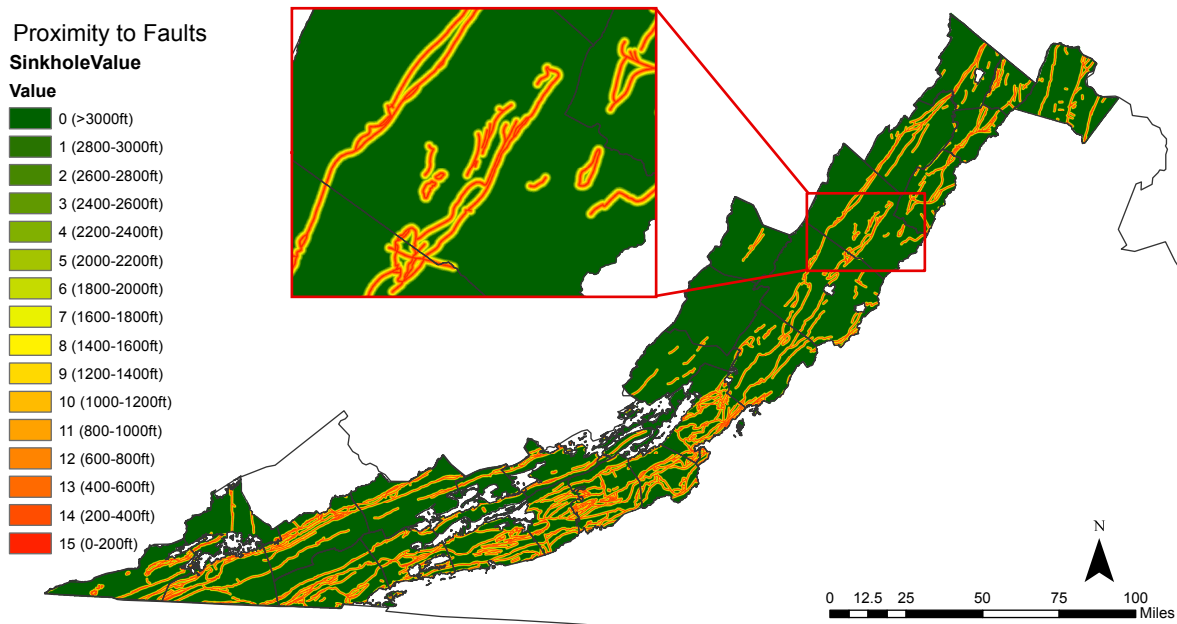
Figure 4: Drainage Class Risk Layer Used in Weighted Combination
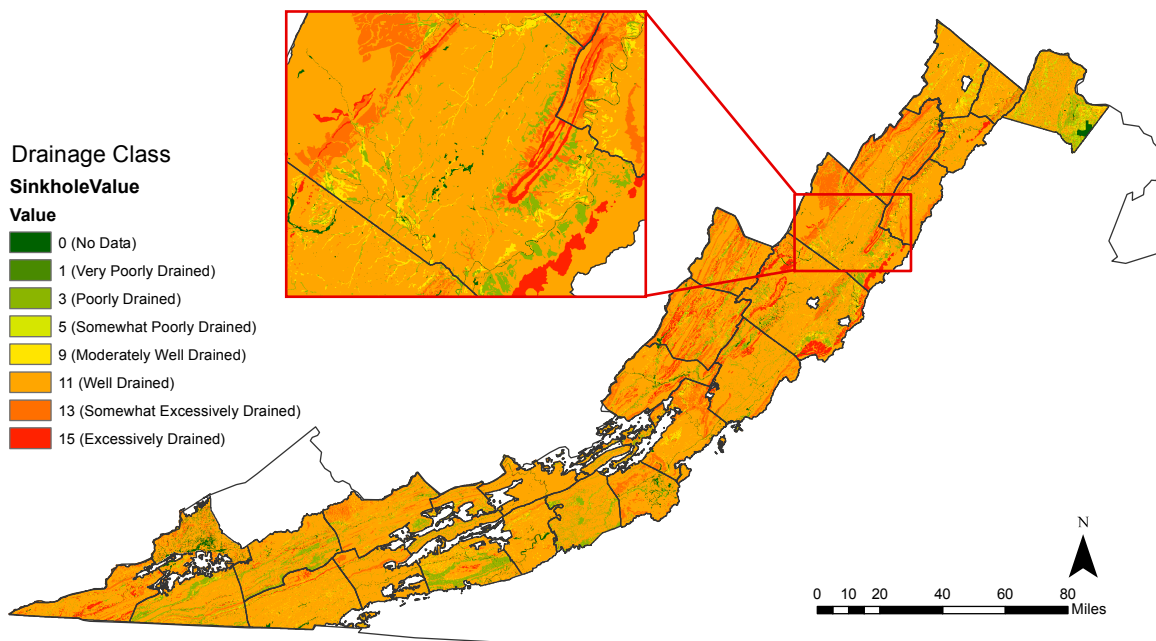
Slope of Incised River Banks

A higher risk for sinkhole formation near incised river banks can be attributed to a higher hydraulic gradient and resulting increase in nearby groundwater flow. To determine the degree of incised rivers, surrounding slopes of riverbanks were identified based on digital elevation model rasters imported from the USGS National Map and extracted to mask the 27 counties of interest. Virginia rivers taken from the US National Weather Service (USNWS) geospatial data were added to the map layer and the elevation models were clipped to only exist within the half mile buffer zones around each river, a region constructed using the "Multiple Ring Buffer" tool in ArcMap. This buffer zone width was constructed because USNWS rivers are linear shapefiles, meaning river width is not taken into consideration in the spatial data. Half a mile was estimated as appropriate in capturing the full river widths and thus riverbank slopes. After surrounding digital elevation models were clipped, the slope of the elevation models were computed using the ArcMap "Slope" tool. The total range of slopes were split into 15 equally incremented categories and slopes of regions were reclassified into one of the 15 consequent categories, assigning low slopes in the regions surrounding rivers to a Sinkhole Value of 1 and high slopes to a Sinkhole Value of 15 (**Figure 5**). Higher slope corresponded to steeper surrounding river banks and a more deeply incised region of the river, defining a higher risk for karst because of a steepened hydraulic gradient and subsequent increase in groundwater flow.

Figure 5: Slope of Incised River Bank Risk Layer Used in Weighted Combination
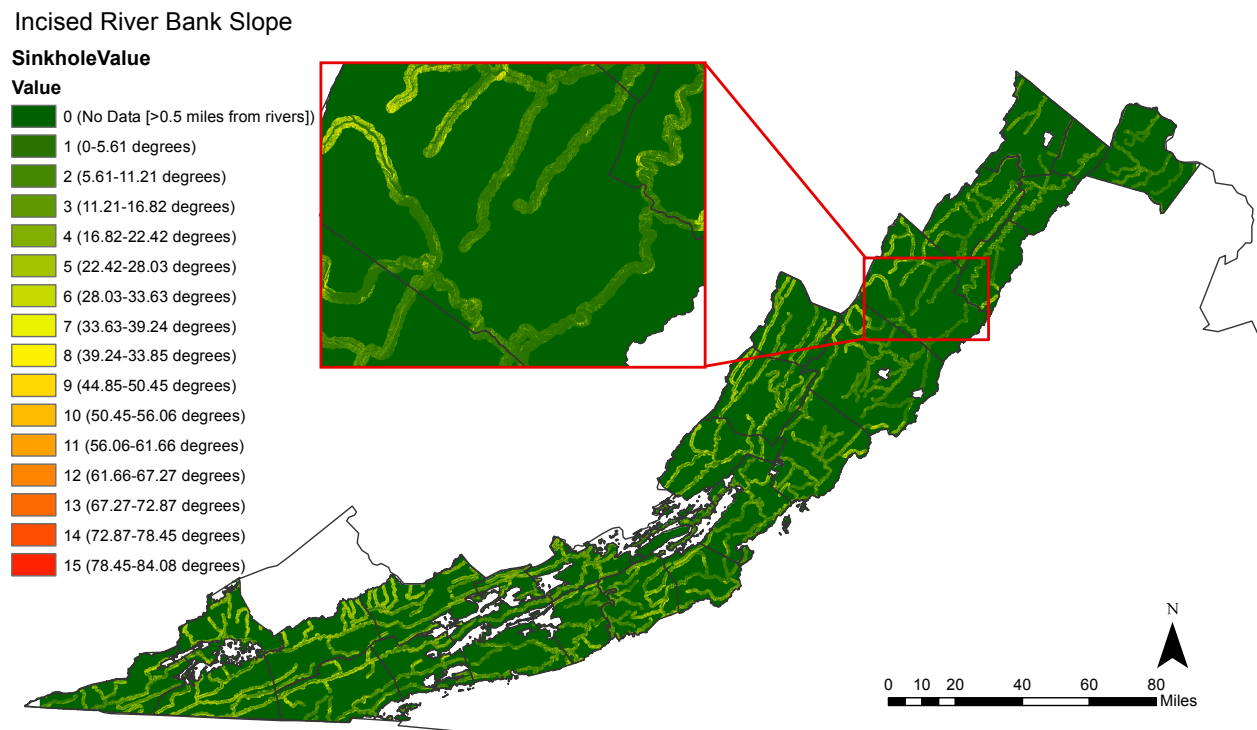
Depth of Overlying Soil

The timescale during which sinkholes may develop is much shorter in regions of shallow depth to bedrock. Thus a sinkhole risk layer representing depth to bedrock layer was created using SSURGO data found in the *Mapunit Aggregated Attribute* (*MUAGGATT)* table. The bedrock minimum depth entry for each map unit along the 27 counties was recorded into a table using a

Python script, based on a common map unit key found in the *Mapunit* and the *MUAGGATT* tables (**Appendix: Figure 13**). The resulting output table was saved into an Access Database and was then imported into ArcMap, where it was spatially joined with the map units. The depth ranges were converted to raster, and were then reclassified into 15 equal intervals, defined for each 15cm increment below the surface. Shallow soil cover overlying the bedrock was assigned the highest Sinkhole Value of 15, and the deepest amount of soil cover was assigned the lowest Sinkhole Value of 1 (**Figure 6**).

Figure 6: Depth of Overlying Soil to Bedrock Risk Layer Used in Weighted Combination
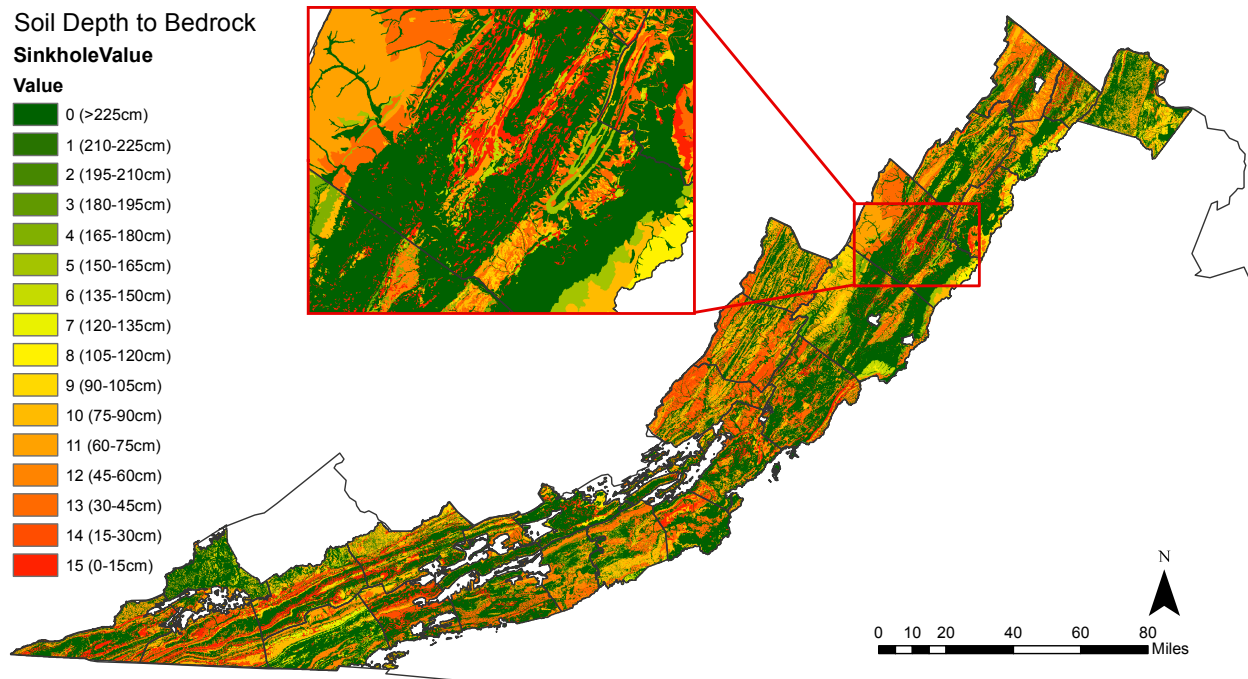
Statistical Analysis (Creating Weights)
In order to analyze how these individual risk layers influenced sinkhole risk in the karst counties of Virginia, a series of maps were created using the ArcMap Raster Calculator. Twenty-eight different risk maps were created by combining the five individual risk layer raster images using a series of different chosen weights. The equation in which the individual risk layers were combined is seen below:

$$
\begin{aligned}
(A \times Bedrock\ Type\ Raster) &+ (B \times Proximity\ to\ Faults\ Raster) \\
&+ (C \times Minimum\ Depth\ to\ Bedrock\ Raster) \\
&+ (D \times Drainage\ Class\ Raster) \\
&+ (E \times Slope\ of\ Incised\ River\ Banks\ Raster) \\
&= Weighted\ Combination\ Risk\ Map
\end{aligned}
$$

where A, B, C, D, and E are the arbitrarily chosen weights assigned to its corresponding combination, and A+B+C+D+E = 1. Because of the infinite number of possibilities to assign the weights, weight values were chosen to be increments of one tenths. Bedrock type has shown in much of the existing karst literature to be the most influential risk factor contributing to sinkhole

formation and should therefore most likely be assigned the highest weight. Therefore combinations were chosen giving bedrock type the highest weight for all combinations except a control combination, combination one, where each layer was assigned equal weight **(Table 2)**.

Table 2: Weighted Combinations of Risk Layers

Combo	Bedrock	Faults	Depth	Drainage	Slope	Combo	Bedrock	Faults	Depth	Drainage	Slope
		Weights						*Weights*			
1	0.2	0.2	0.2	0.2	0.2	15	0.3	0.2	0.1	0.2	0.2
2	0.6	0.1	0.1	0.1	0.1	16	0.3	0.1	0.2	0.2	0.2
3	0.5	0.1	0.1	0.2	0.1	17	0.6	0.2	0	0.1	0.1
4	0.4	0.1	0.2	0.2	0.1	18	0.6	0.2	0.1	0	0.1
5	0.4	0.1	0.2	0.1	0.2	19	0.6	0.2	0.1	0.1	0
6	0.5	0.2	0.1	0.1	0.1	20	0.6	0.1	0.2	0	0.1
7	0.5	0.1	0.2	0.1	0.1	21	0.6	0.1	0.2	0.1	0
8	0.5	0.1	0.1	0.1	0.2	22	0.6	0	0.2	0.1	0.1
9	0.4	0.2	0.2	0.1	0.1	23	0.6	0	0.1	0.2	0.1
10	0.4	0.2	0.1	0.2	0.1	24	0.6	0.1	0	0.2	0.1
11	0.4	0.2	0.1	0.1	0.2	25	0.6	0.1	0.1	0.2	0
12	0.4	0.1	0.1	0.2	0.2	26	0.6	0.1	0.1	0	0.2
13	0.3	0.2	0.2	0.1	0.2	27	0.6	0.1	0	0.1	0.2
14	0.3	0.2	0.2	0.2	0.1	28	0.6	0	0.1	0.1	0.2

Upon completion of the 28 distinctly weighted risk maps, the map that most closely and statistically corresponded to the imported data of mapped sinkhole locations would be chosen. A Python script that could loop through each combination, performing the following steps, was constructed (**Appendix: Figure 14**). Spatial data containing existing mapped sinkholes was added into the map. Each weighted risk map was converted from a raster image to a polygon shapefile based on its Sinkhole Level (1-15) so that it could be clipped into the boundary of the existing sinkholes *(1)* (**Figure 7**). Using the "Dissolve" tool, the newly clipped risk map was condensed into fifteen total zones, based on its defined sinkhole levels. A new field was added to the attribute table of the polygon to calculate the total areas of each distinctive Sinkhole Level. In Excel, the areas of sinkholes corresponding to each Sinkhole Level, the total area of existing sinkholes, and the percentages of each individual level compared to the total sinkhole area were computed and recorded. For simplicity, the 15 Sinkhole Levels were condensed into five risk zones, defining values 1-3 as a Low Risk Zone, 4-6 as a Medium-Low Risk Zone, 7-9 as a Medium Risk Zone, 10-12 as Medium-High Risk Zone, and 13-15 as a High Risk Zone.

Figure 7: Visualization of Methodology in Creating Weights

Combination Polygon   Overlaid Sinkholes   Clipped Combination
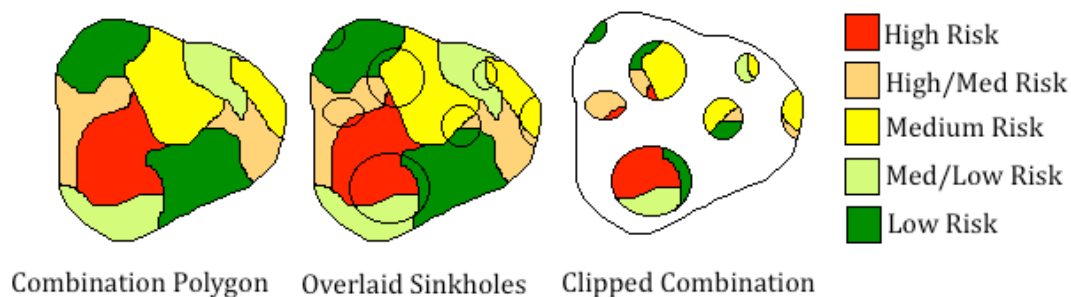
A table was created storing each combinations' ID number and percentages of existing sinkholes areas when compared to the total area of existing sinkholes, per risk zone (**Table 3**). The ideal percentages of existing sinkhole areas per risk zone were defined so they could be statistically compared with the observed percentages using a Residual Sum of Square (RSS) error test. In an ideal situation, there would be no actual sinkholes found in the Low Risk Zone and the highest percentage of actual mapped sinkholes would be found in the High Risk Zone, with a linear relationship between those zones in between and the sum of the total percentages being 100. Hence it is predicted that percentages should be 0, 10, 20, 30, and 40% respectively, per increasing risk zone.

The RSS error test is used to compare the actual and the predicted values of a trend. The term is determined by computing:

$$RSS = \sum (Actual - Predicted)^2$$

Error between the actual model and the ideal model provides an idea of how well the experimental data fits what is expected. Thus the goal is to minimize the RSS. RSS values for each combination were computed and ranked. The combination with the lowest RSS is the value most closely matching the ideal situation and therefore is the combination that will be used for the final map.

Table 3: Percentage of Sinkhole Areas Per Risk Zone of Total Sinkhole Area

Combination	Percent of Total Sinkhole Area Per Category				
	Category				
	1 to 3	4 to 6	7 to 9	10 to 12	13 to 15
1	8.92	45.14	33.52	12.33	0.08
2	13.57	20.22	10.7	43.19	12.32
3	9.24	15	24	45.33	8.78
4	8.37	21.05	35.29	27.21	8.09
5	10.93	27.48	31.77	29.25	0.58
6	11.6	22.45	28.65	29.81	7.5
7	10.22	23.49	29.36	25.63	11.29
8	13.7	23.95	25.6	36.16	0.59
9	9.33	25.77	32.3	25.71	6.9
10	8.59	21.58	35.22	29.46	5.15
11	12.1	26.86	43.78	16.89	0.37
12	9.57	24.33	33.25	32.69	0.15
13	14.18	41.48	29.96	14.01	0.36
14	7.8	25.96	36.49	25	4.76
15	8.66	29.1	47.93	14.23	0.08
16	8.61	28.63	35.01	27.6	0.15
17	13.09	20.58	10.88	44.91	10.54
18	14.7	21.17	22.78	31.95	9.4
19	11.4	19.08	13.3	41.09	15.1

	1 to 3	4 to 6	7 to 9	10 to 12	13 to 15
20	13.26	22.15	22.88	25.23	16.48
21	10.27	20	13.54	26.65	29.54
22	10.52	21.41	12.4	27.09	28.58
23	11.74	18.19	13.6	27.76	28.71
24	12.07	17.44	13.91	43.92	12.66
25	9.42	15.8	17.6	25.12	32.05
26	16.11	22.63	20.36	39.32	1.58
27	16.04	21.55	7.65	54.01	0.75
28	15.17	20.94	8.86	51.83	3.2

RESULTS
Data Analysis
Combination 25 had the smallest RSS when compared with the predicted model (**Table 4, Figure 8**). The percentages of sinkhole areas found in each risk zone for this combination most closely followed the trend of what was predicted. Of the 28 combinations, only four had trends with obvious similarities to the predicted. Ideally, the percentage of sinkhole areas per risk zone would continuously increase linearly between categories. Combinations 21, 22, and 23 increased in percentage of sinkhole areas for three of the four transitions between risk zones, while combination 25 increased continuously across all four transitions (**Figure 8**).

Table 4: Residual Sum of Square Error Per Combination

| | *Category* | | | | | |
Combination	*1 to 3*	*4 to 6*	*7 to 9*	*10 to 12*	*13 to 15*	*RSS Error*
1	8.92	45.14	33.52	12.33	0.08	3403.0117
2	13.57	20.22	10.7	43.19	12.32	1315.2418
3	9.24	15	24	45.33	8.78	1336.0749
4	8.37	21.05	35.29	27.21	8.09	1451.9757
5	10.93	27.48	31.77	29.25	0.58	2118.0471
6	11.6	22.45	28.65	29.81	7.5	1420.6711
7	10.22	23.49	29.36	25.63	11.29	1217.3991
8	13.7	23.95	25.6	36.16	0.59	2004.7462
9	9.33	25.77	32.3	25.71	6.9	1601.0459
10	8.59	21.58	35.22	29.46	5.15	1654.347
11	12.1	26.86	43.78	16.89	0.37	2738.567
12	9.57	24.33	33.25	32.69	0.15	2067.7549
13	14.18	41.48	29.96	14.01	0.36	3118.2741
14	7.8	25.96	36.49	25	4.76	1854.3393
15	8.66	29.1	47.93	14.23	0.08	3062.1898
16	8.61	28.63	35.01	27.6	0.15	2240.2916
17	13.09	20.58	10.88	44.91	10.54	1456.6586
18	14.7	21.17	22.78	31.95	9.4	1288.7498
19	11.4	19.08	13.3	41.09	15.1	1000.2945

20	13.26	22.15	22.88	25.23	16.48	907.6878
21	10.27	20	13.54	26.65	29.54	367.8386
22	10.52	21.41	12.4	27.09	28.58	437.503
23	11.74	18.19	13.6	27.76	28.71	378.3454
24	12.07	17.44	13.91	43.92	12.66	1179.3686
25	9.42	15.8	17.6	25.12	32.05	215.1533
26	16.11	22.63	20.36	39.32	1.58	1982.1374
27	16.04	21.55	7.65	54.01	0.75	2660.2492
28	15.17	20.94	8.86	51.83	3.2	2304.701

Figure 8: Sinkhole Area Percentages Per Risk Zone of Total Sinkhole Area
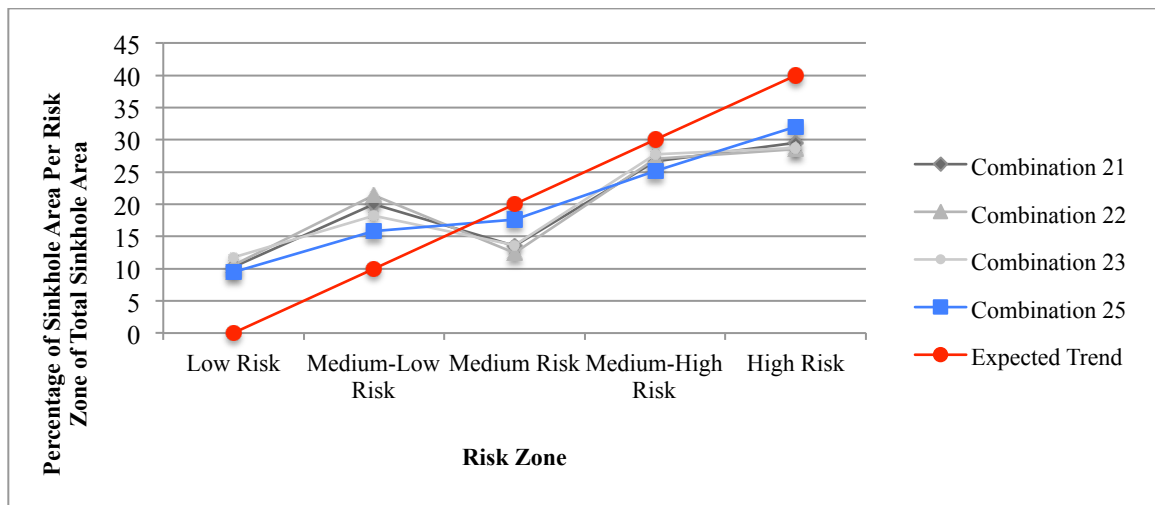
Final Map
The final map, created from combination 25, was created using the following equation:

$$Weighted\ Map = (0.6 \times BedrockLayer) + (0.1 \times ProximityToFaultsLayer) + (0.2 \times DrainageClass) + (0 \times SlopeOfIncisedRiverBanks) + (0.1 \times DepthToBedrockLayer)$$

The final risk shown for all 27 counties is shown in **Figure 9**.

DISCUSSION
Combination Twenty-Nine
Following the statistical analysis perfomed on the 28 map combinations during which the best representation for karst risk was identified, one final combination was created based on patterns observed in the four combinations with the smallest RSS error (**Table 4**) and whose percentage of sinkhole areas increased for at least three of the four transitions between risk zones: combinations 21, 22, 23, and 25. New weights were assigned to each risk layer determined by the weight similarities between the four previous combinations (**Table 5**). The steps taken in the previous statistical analysis were repeated to determine if this new combination would better represent the sinkhole risk than combination 25. However, the RSS error of combination 29 was greater than the RSS error of combination 25 (**Figure 5**), and so combination 25 remained the

best option for the final sinkhole risk map.

Table 5: Patterns in Best Combinations' Weight Assignments Ranked by Ascending RSS Error and New Weight Assignment for Combination 29

Combo	Bedrock	Faults	Depth	Drainage	Slope	RSS Error
25	0.6	0.1	0.1	0.2	0	215.1533
21	0.6	0.1	0.2	0.1	0	367.8386
23	0.6	0	0.1	0.2	0.1	378.3454
22	0.6	0	0.2	0.1	0.1	437.503
29	0.6	0.067	0.15	0.15	0.033	316.5727

Bedrock Type
The final weighted map equation assigned bedrock a weight of 0.6, telling us that bedrock has the most significant impact on predicting sinkhole risk and confirming what was expected. This assigned weight of 0.6 for bedrock type remained consistent across the four combinations with the smallest RSS error, confirming that it should represent the most influential risk factor on the final map. Karst terrain is a direct result of water chemically reacting with limestone bedrock, hence type unquestionably influences the likeliness of a sinkhole prone region.

Proximity to Fault Lines
Proximity to faults was assigned a weight of 0.1, meaning it plays a minimal role in determining sinkhole risk of a region. Because certain combinations excluded this risk layer entirely, such as combinations 22 and 23, and yet combination 26 was determined to be statistically more significant than those, proximity to faults has proven to affect the risk map and thus play a role in determining sinkhole risk regions in Virginia. However it is far less of a factor than bedrock type.

Drainage Class
Drainage class was given a weight of 0.2 and thus it represents a more significant risk factor in determining sinkhole risk of a region than proximity to fault lines, slope of incised river banks, or soil depth to bedrock. A higher drainage class obviously influences the results, yet its assigned weight when compared to bedrock type show that it still should not stand alone in determining regions at risk for sinkhole formation.

Slope of Incised River Banks
The risk layer combination that most closely matched the ideal sinkhole area percentages gave no weight to the incised river banks layer. This means that the incised river bank layer most likely did not influence sinkhole risk. While there remains the possibility that it could play a minute role in predicting sinkhole risk, out of the 28 combinations created, the best option eliminated this layer entirely.

Depth of Overlying Soil
Depth to overlying soil was assigned a weight of 0.1, and so, like the proximity to fault layer, minimum soil depth to bedrock plays an existing yet nearly insignificant role in predicting sinkhole risk in Virginia.

Figure 9: Final Sinkhole Risk Map with Interstates Overlain
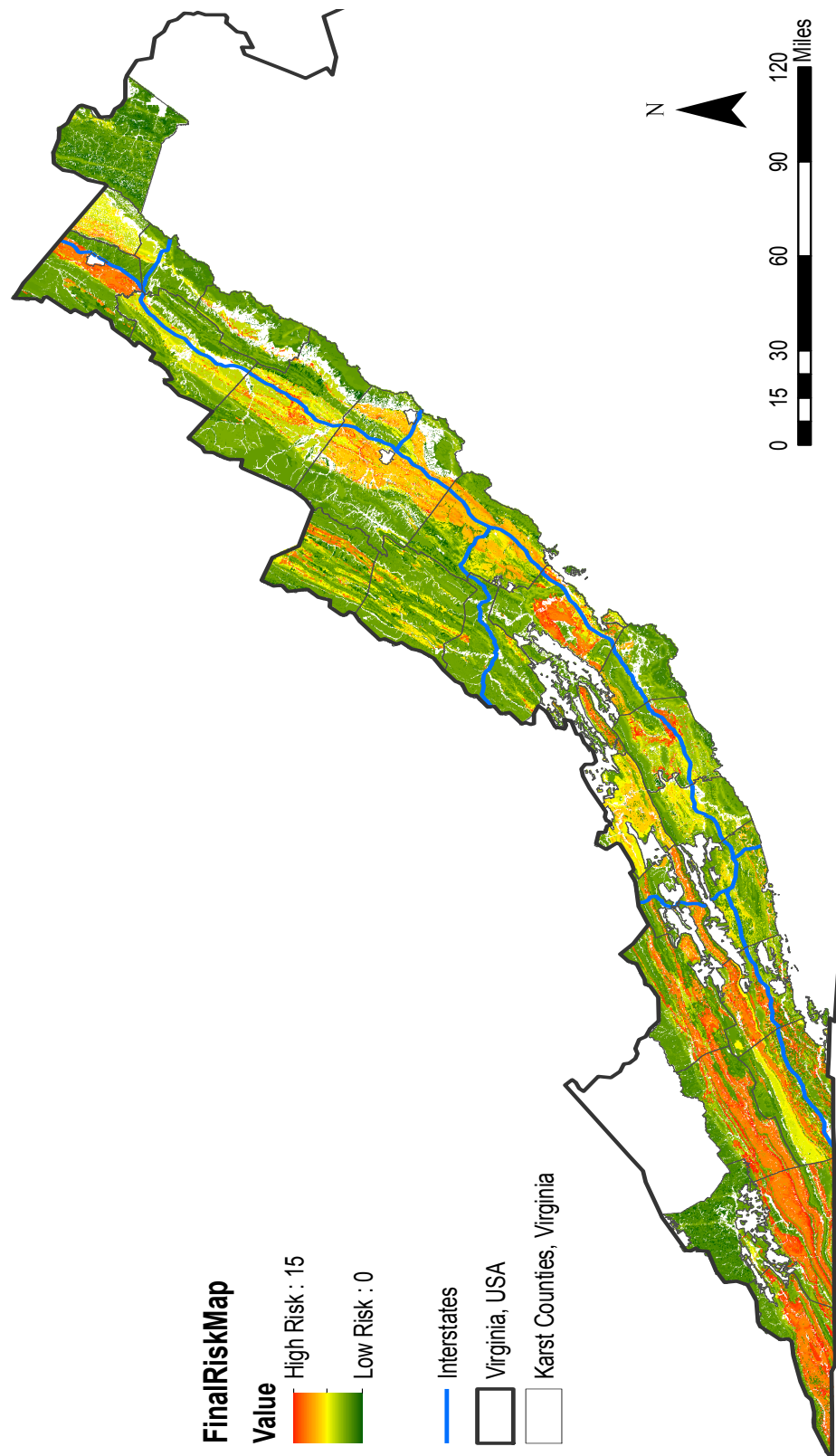
Interpreting the Final Map
Displaying the current USGS Karst Terrain map beside the final sinkhole risk map allows us to contemplate if the final risk map makes sense **(Figure 10)**. Since karst terrain is defined to be terrain containing subsurface fissures, caverns, and voids resulting from chemically dissolved limestone bedrock, if the two maps relate, then we can be confident in the conclusions reached. From a visual comparison, it can be established that the resulting values make sense, since the regions with karst terrain on the USGS map align with the higher risk regions on the sinkhole risk map.

Sources of Error
While these results serve as a general guideline and a good start to mapping karst regions in Virginia, it is important to understand that a risk map created based on generalities cannot be substituted for a site-specific analysis. Different karst landforms relate to one another, but the combinations and behaviors of the relationship between local, hydrological, and climactic conditions are infinite *(20)*. While this map may provide a general understanding of the karst terrain in Virginia, the final product and the methods taken to reach it are specific to the region of interest. It must be noted that one karst technology applied to a specific region cannot necessarily be used in a different karst region. Such assumptions introduce a large problem when using a GIS for risk analysis – it becomes entirely too easy to use techniques learned from one study and apply them elsewhere when it may not be applicable.

The final weights used to combine the layers into the ultimate sinkhole risk map were based on a statistical comparison the constructed predicted risk regions with existing sinkholes mapped by David Hubbard. However, to make clear that his sinkholes were mapped as a guideline and not a set of perfectly defined structures, he digitized his data using a scale that was 10 times less accurate than the scale the public had hoped he would. Furthermore, the aerial photography used in Hubbard's study cannot accurately detect all sinkholes due to aircraft tilt, which creates the appearance of a sinkhole where it may not exist and does not recognize shallower sinkholes at all. This proves just how rough Hubbard's representation of sinkhole location is and could serve as an important source of error in choosing the appropriate weights for the combination of risk layers.

By using the USDA Soil Survey data, a degree of error was inevitable due to the fact that the data tables from which a significant portion of the risk layers were derived were incomplete for specific factors in a majority of the counties being focused on. Additionally, the SSURGO data contained data values corresponding to Virginia *counties* but not to major Virginia cities, who likely have their own GIS database. This led to voids in the final map, denoting cities with no representative SSURGO data.

The type of data available further limited the scope of this study. The slope of incised riverbanks risk layer was determined using digital elevation models, raster images created using remote sensing. While these are helpful for analyzing the surface topography of the region, difficulties in karst analyses arise since the directions of subsurface flow and hydrologic connections between sinkholes or subsurface aquifers are not represented in them *(21)*. It would be much more useful to know the direction of water flow in those surrounding incised river banks, since large slopes on the surface do not necessarily denote fast paced water travel through subsurface fissures and waterways.

Figure 10: Comparison of USGS Karst Terrain with Final Risk Map
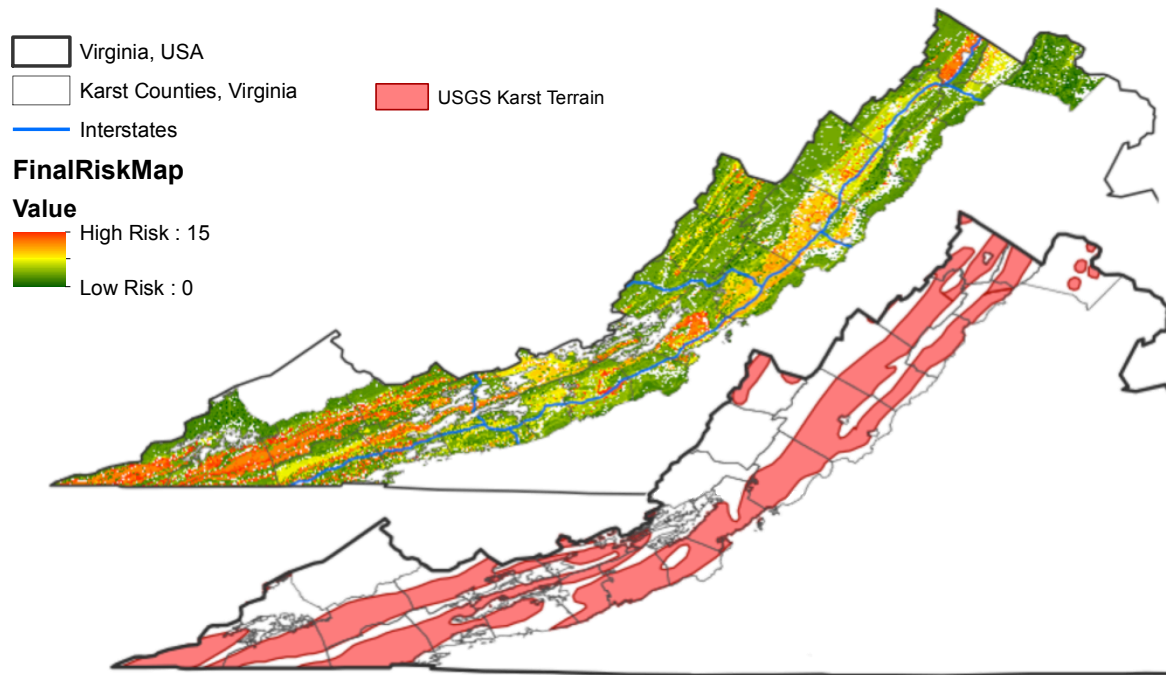
Source: United States Geologic Survey, provided by Virginia Department of Mines, Minerals, and Energy

Finally, the risk layers compiled in creating the final map were created and assigned sinkhole values based on hypotheses derived from background knowledge. Sinkhole values were determined using engineering judgment and personal interpretation. Therefore it is always possible that a sinkhole value was assigned incorrectly. For example, drainage class assumed that high drainage regions had the highest risk for sinkhole formation, because it was predicted that water disappeared, rather than puddling, into the ground through karst caverns and fissures. Therefore, the drained water would assist in chemically dissolving surrounding cavern walls expanding them and creating sinkholes beneath the earth's surface. However high drainage could also be interpreted as water flowing down surface slopes into existing bodies of water or rivers, and not into underground fissures. Similarly, the proximity to fault risk factor was chosen based on background knowledge of sinkholes and karst terrain. However, faults usually influence the formation of sinkholes only when they are active and changing. Virginia faults have, for the most part, been inactive for hundreds of years, and therefore could perhaps have no present influence on sinkhole formation.

CONCLUSION

This analysis used a geographic information system to create a map that best represents regions most at risk of sinkhole formation in the karst counties of Virginia. The resulting map can be used as a recommendation for VDOT in determining the roadways and regions most susceptible to experiencing sinkholes in an attempt to reduce reported sinkhole damage to life and property. Five risk factor layers – bedrock type, proximity to fault lines, drainage class of a region, slope of incised river banks, and soil depth to bedrock – were combined in an equation based on

statistically tested weight assignments for each layer. Resulting combination maps were compared with existing mapped sinkholes to statistically verify which combination most closely fit the expected results and which regions along Virginia interstates and heavily trafficked roadways may require immediate safety evaluations. In utilizing these maps, VDOT can pinpoint areas of roadway and infrastructure that have potentially high risk of subsidence and use this information for roadway repair and planning purposes (**Figure 11**). With a goal of holding paramount the safety of the people, the environment, and their wellbeing, VDOT can be aware of regions with high sinkhole risk and be conscientious in their work to be certain that, in the case of a sinkhole forming, no persons or property will be damaged.

From the analysis conducted in this study, it was found that sinkhole risk was most sensitive to the type of bedrock underlying the region, with pure limestone and combinations of limestones and dolomites defining risk more so than regions of partial carbonate bedrock or entirely clastic (and non-carbonate) bedrock. Drainage class, in particular regions of high drainage over those of low drainage, played a higher role in the risk prediction than any other remaining layer, but it remained significantly less influential than bedrock type. Proximity to faults and minimal depth to bedrock only slightly affected in the projection of where sinkholes might form, but their existence in the weighted equation of layers was more representative of existing mapped sinkholes than their absence. However, steep slopes of incised river banks in the karst regions of Virginia played no role in the final map allowing that risk layer to be eliminated from the final weighted combination map.

Figure 11: Final Map Displaying Examples of High, Medium, and Low Risk Near Interstates
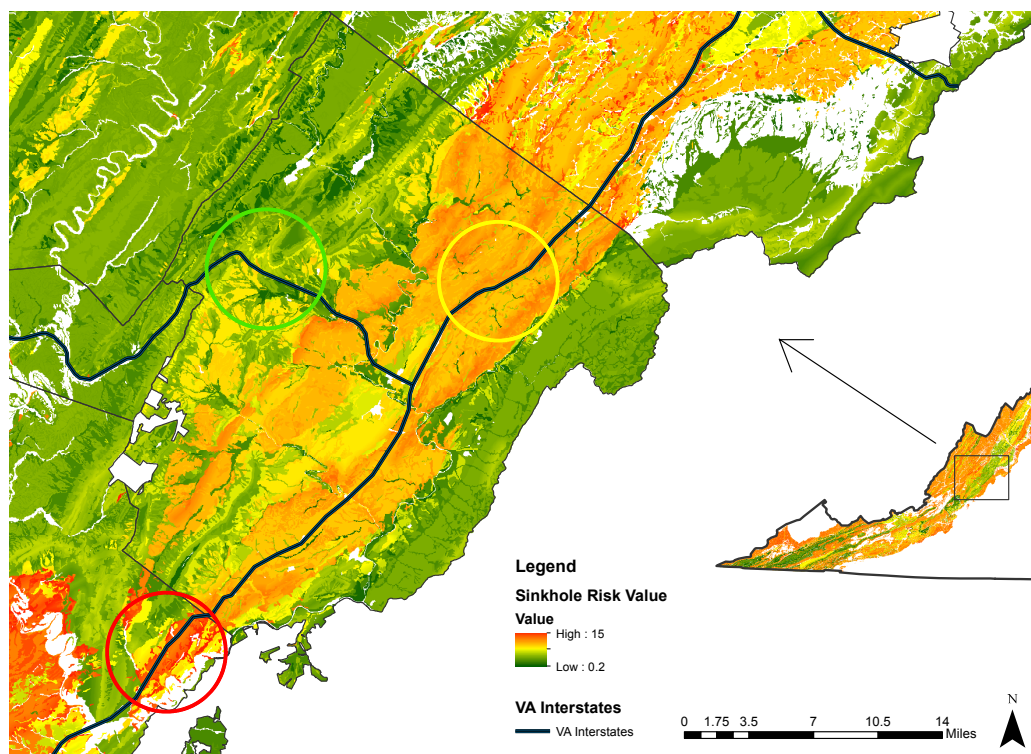
Recommendations for Future Work

The most accurate method of determining a risk map for karst terrain in the state of Virginia would be field mapping across the region, however the investigation would be time consuming and expensive. To improve the results, introducing more risk layers would allow the resultant map to be most accurate. Five layers were used as a basis for this kind of risk map but increasing the influencing layers would allow for an even more correlated map when compared with mapped existing sinkholes. The mapping process could also be further improved if more GIS data was available. The SSURGO database did not have tabular data representing all risk factors for the region of interest and some potentially useful factors could not be contributed to the analysis, such as bathymetry of rivers and streams. Similarly, bedrock type was determined based on the data available without taking into consideration the depth at which that bedrock exists. If this kind of data was recorded and available, it would have been more useful to find a specific horizon beneath the surface and remain consistent in mapping only the bedrock type across that depth, and a better depiction of sinkhole risk in the region of interest could be drawn.

The data represented in the SSURGO database is not detailed enough to be a trustworthy and entirely accurate source for soil data across such a vast region. It acts as a general guideline for soil types yet does not have high enough resolution to represent the highly variable soil patterns of the karst counties of Virginia *(4)*. By using higher resolution data, the resulting risk map could be extremely enhanced. Higher resolution data can be captured using remote sensing techniques such as Light Detection and Ranging (LiDAR), for example, where more accurate elevation models can be created by measuring properties of scattered light to find the range or other information about a distant target *(22)*. LiDAR would be helpful in creating a map of existing sinkholes to compare with a final risk map, because it would eliminate most possibilities of inaccuracies found in mapping sinkholes using aerial photography. It is impossible to detect all sinkholes by aerial photography since aircraft tilt makes it difficult to recognize sinkholes if the subsidence is too shallow, and often, illusions lead to detected sinkholes that do not really exist. It must be noted that, similar to aerial photography, LiDAR cannot provide a fully representative elevation map in tree covered areas, and so field checks will be required in those regions. However it is an appropriate method to use for the most accurate mapped sinkhole representation, rather than aerial photography, since Hubbard estimates it will take approximately 250 years to map every single Virginia sinkhole using only aerial photography and following field checks.

LITERATURE CITED

(1) Hubbard, Jr. D.A. Sinkhole Distribution of the Valley and Ridge Province, Virginia. In *Geotechnical and Environmental Applications of Karst Geology and Hydrology: Proceedings of the Eighth Multidisciplinary Conference on Sinkholes*, April 2001, pp.33-36.

(2) Hubbard, Jr. D.A. Selected Karst Features of the Central Valley and Ridge Province, Virginia. *Virginia Division of Mineral Resources*, Publication 83, 1988.

(3) The USGS Water Science School. *Sinkholes,* July 2014, http://water.usgs.gov/edu/sinkholes.html. Accessed August 2014.

(4) Hyland, Sara E. Analysis of Sinkhole Susceptibility and Karst Distribution in the Northern Shenandoah Valley, VA: Implications for Low Impact Development (LID) Site Suitability

Model. Master Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, May 2005.

(5) Sowers, G.F. *Building on Sinkholes: Design and Construction of Foundations in Karst Terrain.* American Society of Civil Engineers Press, New York, 1996.

(6) Belo, B.P. Natural Hazard Mitigation Planning for Karst Terrains in Virginia. Master Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, May 2003.

(7) William and Mary Department of Geology. *Valley and Ridge Province, Geology of Virginia,* http://web.wm.edu/geology/virginia/provinces/valleyridge/valley_ridge.html. Accessed July 2014.

(8) Ivey Burden, L. Karst Topography: Noninvasive Geophysical Detection Methods and Construction Techniques. Virginia Council for Transportation Research and Innovation. Final Contract Report. Project UVA-2012-01, 2013, pp.1-22.

(9) Dai, J., Lei, M., Lui, W., Tang, S., and S. Lai. An Assessment of Karst Collapse Hazards in Guilin, Guangxi Province, China. In *Sinkholes and the Engineering and Environmental Impacts of Karst*, 2008, pp.156-164.

(10) Whitman, D. and T.L. Gubbels. Applications of GIS Technology to the Triggering Phenomena of Sinkholes in Central Florida. In: Beck, B.F., Pettit, A.J., and Herring, J.G., eds., *Hydrogeology and Engineering Geology of Sinkholes and Karst*, 1999, pp.67-73.

(11) *Commonwealth of Virginia Hazard Mitigation Plan.* Section 3-14. Virginia Department of Emergency Management, Virginia Tech Center for Geospatial Information Technology, 2003.

(12) Waltham, A. and P. Fookes. Engineering Classification of Karst Ground Conditions. In *Quarterly Journal of Engineering Geology and Hydrogeology*, Volume 36, 2003, pp.101-118.

(13) Moore, H. and L. McDowell. The Development and Use of Karst Maps in the Location of Highways in Eastern Tennessee. In *Sinkholes and the Engineering and Environmental Impacts of Karst*, 2008, pp.680-693.

(14) Muckel, Gary B., ed. *Understanding Soil Risks and Hazards: Using Soil Survey to Identify Areas with Risks and Hazards to Human Life and Property.* United States Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center, 2004.

(15) Hubbard, Jr. D.A. Selected Karst Features of the Northern Valley and Ridge Province, Virginia. In *Virginia Division of Mineral Resources*, Publication 44, 1983.

(16) Green, J.A., Marken, W.J., Alexander, Jr., E.C. and S.C. Alexander. Karst Unit Mapping using Geographic Information System Technology, Mower County, Minnesota, USA. In *Environmental Geology*, Volume 42, Issue 5, 2002, pp.457-461.

(17) Doctor, D.H., Weary, D.J., Orndorff, R.C., Harlow, Jr. G.E., Kozar, M.D., and D.L. Nelms. Bedrock Structural Controls on the Occurrence of Sinkholes and Springs in the Northern Great Valley Karst, Virginia and West Virginia. In *Sinkholes and the Engineering and Environmental Impacts of Karst*, Issue 183, 2008, pp.12-22.

(18) Doctor, K.Z., Doctor, D.H., Kronenfeld, B., Wong, D.W., and D.K. Brezinski. Predicting Sinkhole Susceptibility in Frederick Valley, Maryland, Using Geographically Weighted Regression. In *Sinkholes and the Engineering and Environmental Impacts of Karst*, 2008, pp.243-256.

(19) Zhou, W., Beck, B.F., and A.L. Adams. Application of Matrix Analysis in Delineating Sinkhole Risk Areas Along Highways (I-70 near Frederick, Maryland). In *Environmental Geology*, Issue 44, 2003, pp.834-842.

(20) Ford, D.C. and P.F. Williams. *Karst Geomorphology and Hydrology*. Unwin Hyman, London, 1989.

(21) Taylor, C.J., Kaiser, W P., & H.L. Nelson, Jr. (2008). Application of Geographic Information System (GIS) Hydrologic Data Models to Karst Terrain. In *Sinkholes and the Engineering and Environmental Impacts of Karst*, 2008, pp.146-155.

(22) Sangam, Sanu. Light Detection and Ranging. CMR Institute of Technology, No. 132, AECS Layout, Bangalore.

APPENDIX

Figure 11: Python Code Producing Sinkhole Values based on Bedrock Type

```
 1 # -*- coding: utf-8 -*-
 2 # -------------------------------------------------------------------
 3 # 2014_5_16_Python_Weights.py
 4 # Created on: 2014-05-16 4:30:51.00000
 5 #    (generated by ArcGIS/ModelBuilder)
 6 # Description:
 7 #    Assigns weighted sinkhole values based on bedrock type and creates a new
 8 #    table to store them in so they can be manually combined with the mapunit
 9 #    shapefile in ArcMap
10 # Author:
11 #    Alexandra L. Todd
12 # -------------------------------------------------------------------
13
14 # Import python odbc library
15 import pyodbc
16
17 # Save output to .txt table
18 file = open(r'''H:\Research_Thesis\THESIS\Python_Scripts\\Python_Output.txt''' \
19     ,"w")
20
21 file.write("mukey; pmkind; pmorigin; comppct_r; compname; majcompflag; \
22     pmgroupname; SH_Val; CountRepeats; Weight_Calc; Percent_Sum; unique_mukey; \
23     Weight_Final \n")
24
25 # Define counties of interest
26 counties = ['105','195','169','167','191','173','197','021','155','071',\
27     '121','185','045','161','023','005','017','091','163','015','165','139',\
28     '171','187','069','043','107']
29
30 for county in counties:
31     print('Working on county_' + county)
```

```python
33      # Connect to database
34      db_file = r'''H:\\Research_Thesis\\THESIS\\Co_Databases\\soildb_US_\
35      2003_''' + county + r'''.mdb'''
36      user = ''
37      password = ''
38      conn_str = 'DRIVER={Microsoft Access Driver (*.mdb)};DBQ=%s;UID=%s;PWD=%s' %
39                      (db_file, user, password)
40      conn = pyodbc.connect(conn_str)
41      cursor = conn.cursor()
42
43      # Create Component_Table1
44      Make_Table = """ (SELECT mapunit.mukey,copmgrp.cokey, copmgrp.copmgrpkey, \
45      copm.pmkind, copm.pmorigin, component.comppct_r,
46      component.compname, component.majcompflag, copmgrp.pmgroupname
47      INTO Comp_Table FROM ((((mapunit INNER JOIN component ON (mapunit.mukey = \
48      component.mukey))
49      INNER JOIN copmgrp ON (component.cokey = copmgrp.cokey))
50      INNER JOIN copm ON (copmgrp.copmgrpkey = copm.copmgrpkey))
51      INNER JOIN chorizon ON (component.cokey = chorizon.cokey))) """
52      cursor.execute(Make_Table)
53      print('done with add/join in county_' + county)
54
55      # Add ID column to table
56      cursor.execute("ALTER TABLE Comp_Table ADD ID autoincrement(1,1)")
57      print('done with adding Sinkhole_Value field in county_' + county)
59      # Delete duplicate entries in table
60      Delete_Duplicates = """ DELETE Comp_Table.*
61      FROM Comp_Table
62      WHERE ((((SELECT COUNT (*) FROM Comp_Table AS L
63      WHERE (L.cokey & L.mukey & L.copmgrpkey & L.pmkind & L.pmorigin & \
64      L.comppct_r = Comp_Table.cokey & Comp_Table.mukey & Comp_Table.copmgrpkey \
65      & Comp_Table.pmkind & Comp_Table.pmorigin & Comp_Table.comppct_r)
66      AND (L.ID <= Comp_Table.ID)))>1)) """
67      cursor.execute(Delete_Duplicates)
68      print('done with deleting duplicates in county_' + county)
69
70      # delete any empty pmorigin
71      cursor.execute("DELETE from Comp_Table WHERE pmorigin IS null")
72      print('done removing all empty origin fields in county_' + county)
73
74      # delete all pmkind = "alluvium"
75      cursor.execute("DELETE from Comp_Table where pmkind = ?", 'alluvium')
76      print('done removing all values where pmkind = "alluvium" in county_' + \
77          county)
78
79      # add sinkhole_value field
80      cursor.execute('ALTER TABLE Comp_Table ADD SH_Val number')
81      print('done with adding sinkhole value in county_' + county)
82
83      # update sinkhole_value
84      update_SV = """ UPDATE Comp_Table SET SH_Val = ? WHERE pmorigin = ? """
85      cursor.execute(update_SV, 4, 'limestone')
86      cursor.execute(update_SV, 4, 'cherty limestone')
87      cursor.execute(update_SV, 3, 'limestone and dolomite')
88      cursor.execute(update_SV, 2, 'limestone and sandstone')
89      cursor.execute(update_SV, 2, 'calcareous siltstone')
```

```
 90      cursor.execute(update_SV, 2, 'calcareous shale')
 91      cursor.execute(update_SV, 2, 'limestone and shale')
 92      cursor.execute(update_SV, 2, 'limestone, sandstone, and shale')
 93      cursor.execute(update_SV, 1, 'granulite')
 94      cursor.execute(update_SV, 1, 'acid shale')
 95      cursor.execute(update_SV, 1, 'dolomite')
 96      cursor.execute(update_SV, 1, 'granite')
 97      cursor.execute(update_SV, 1, 'greenstone')
 98      cursor.execute(update_SV, 1, 'metaquartzite')
 99      cursor.execute(update_SV, 1, 'metasedimentary rock')
100      cursor.execute(update_SV, 1, 'noncalcareous shale')
101      cursor.execute(update_SV, 1, 'noncalcareous siltstone')
102      cursor.execute(update_SV, 1, 'phyllite')
103      cursor.execute(update_SV, 1, 'quartzite')
104      cursor.execute(update_SV, 1, 'sandstone')
105      cursor.execute(update_SV, 1, 'sandstone and shale')
106      cursor.execute(update_SV, 1, 'shale')
107      cursor.execute(update_SV, 1, 'shale and siltstone')
108      cursor.execute(update_SV, 1, 'siltstone')
109      cursor.execute(update_SV, 1, 'sedimentary, unspecified')
110      cursor.execute(update_SV, 1, 'metamorphic and sedimentary rock')
111      cursor.execute(update_SV, 1, 'sandstone and siltstone')
112      cursor.execute(update_SV, 1, 'coal')
113      cursor.execute(update_SV, 1, 'rhyolite')
114      cursor.execute(update_SV, 1, 'graywacke')
115      cursor.execute(update_SV, 1, 'granite and gneiss')
116      cursor.execute(update_SV, 1, 'slate')
117      cursor.execute(update_SV, 2, 'chert')
118      cursor.execute(update_SV, 1, 'igneous and metamorphic rock')
120      cursor.execute(update_SV, 2, 'limestone and siltstone')
121      cursor.execute(update_SV, 1, 'Igneous, metamorphic and sedimentary rock')
122      cursor.execute(update_SV, 1, 'schist')
123      cursor.execute(update_SV, 1, 'gabbro')
124      cursor.execute(update_SV, 1, 'metamorphic rock')
125      cursor.execute(update_SV, 1, 'granodiorite')
126      cursor.execute(update_SV, 1, 'basalt')
127      cursor.execute(update_SV, 1, 'hornfels')
128      cursor.execute(update_SV, 1, 'diabase')
129      cursor.execute(update_SV, 1, 'mica schist')
130      cursor.execute(update_SV, 2, 'calcareous conglomerate')
131      cursor.execute(update_SV, 1, 'conglomerate')
132      cursor.execute(update_SV, 1, 'gneiss')
133      print('done updating Sinkhole_Values in county_' + county)
134
135      # Create a new table "Count_Weights" with counts of mukeys
136      #(repeated mukeys will be used in the weighted equation)
137      Repeat_MU = """ SELECT Comp_Table.*,
138          (SELECT COUNT (*) FROM Comp_Table AS M
139          WHERE (M.mukey =Comp_Table.mukey)
140          AND (M.ID <= Comp_Table.ID)) AS [Counter]
141          INTO Count_Weights FROM Comp_Table """
142      cursor.execute(Repeat_MU)
143      print('done creating new count row in county_' + county)
144
145      # Add two rows for Weighted Sinkhole Value to perform final step in Excel
146      cursor.execute("""ALTER TABLE Count_Weights ADD SH_Weight1 text""")
147      cursor.execute("""ALTER TABLE Count_Weights ADD SH_Weight text""")
```

```
149       # write out all rows remaining into test.csv
150       for row in cursor.execute("SELECT * from Count_Weights"):
151           file.write(str(row[0]) + ";" + str(row[3]) + ";" + str(row[4]) + ";" +\
152           str(row[5]) + ";" + str(row[6]) + ";" + str(row[7])+ ";" + str(row[8]) \
153           + ";" + str(row[10]) + ";" + str(row[11]) + "\n")
154
155 # Close file containing output
156 file.close()
157
158 print('Done with creating table')
```

Figure 12: Python Code Producing Sinkhole Values Based on Drainage Class

```python
 1 # -*- coding: utf-8 -*-
 2 # --------------------------------------------------------------------------
 3 # 2014_7_29_Python_DrainageClass.py
 4 # Created on: 2014-07-29 2:42:53.00000
 5 #   (generated by ArcGIS/ModelBuilder)
 6 # Description:
 7 #   Creates a table with drainage class for all karst counties in Virginia to
 8 #   then manually combine with the mapunit shapefile in ArcMap
 9 # Author:
10 #   Alexandra L. Todd
11 # --------------------------------------------------------------------------
12
13 # import python odbc library
14 import pyodbc
15
16 # Save output to .txt table
17 file = open(r'''H:\Research_Thesis\THESIS\Python_Scripts\\Python_DrainageClass\
18     _Output.txt''',"w")
19
20 file.write("mukey; drainage_class \n")
21
22 # Define counties of interest
23 counties = ['105','195','169','167','191','173','197','021','155','071','121',\
24     '185','045','161','023','005','017','091','163','015','165','139','171',\
25     '187','069','043','107']
26
27 # Loop through counties
28 for county in counties:
29     print('Working on county_' + county)
30
```

```python
31      # Connect to database
32      db_file = r'''H:\\Research_Thesis\\THESIS\\Co_Databases\\soildb_US_\
33          2003_''' + county + r'''.mdb'''
34      user = ''
35      password = ''
36      conn_str = 'DRIVER={Microsoft Access Driver (*.mdb)};DBQ=%s;UID=%s;PWD=%s' %
37                      (db_file, user, password)
38      conn = pyodbc.connect(conn_str)
39      cursor = conn.cursor()
40
41      # Create Component_Table1
42      Make_Table1 = """ (SELECT mapunit.mukey, component.drainagecl
43      INTO Comp_Table2 FROM (mapunit INNER JOIN component ON (mapunit.mukey = \
44          component.mukey))) """
45      cursor.execute(Make_Table1)
46
47      # delete NULL values for drainagecl
48      cursor.execute("DELETE from Comp_Table2 WHERE drainagecl IS null")
49      print('done removing Null drainage class values')
50
51      # write out all rows remaining into Python_DrainageClass_Output.txt
52      for row in cursor.execute("SELECT * from Comp_Table2"):
53          file.write(str(row[0]) + ";" + str(row[1]) + "\n")
54
55 # Close file containing output
56 file.close()
57
58 print('Done with creating table')
```

Figure 13: Python Code Producing Sinkhole Values Based on Minimum Depth to Bedrock

```
 1  # -*- coding: utf-8 -*-
 2  # -------------------------------------------------------------------
 3  # 2014_7_21_Python_DepthToBedrock.py
 4  # Created on: 2014-07-21 11:01:14.00000
 5  #    (generated by ArcGIS/ModelBuilder)
 6  # Description:
 7  #    Creates a table with minimum depth to bedrock for all karst counties in
 8  #    Virginia to then manually combine with the mapunit shapefile in ArcMap
 9  # Author:
10  #    Alexandra L. Todd
11  # -------------------------------------------------------------------
12
13  # Import python odbc library
14  import pyodbc
15
16  # Save output to .txt table
17  file = open(r'''H:\Research_Thesis\THESIS\Python_Scripts\\Python_BedrockDepth\
18      _Output.txt''',"w")
19
20  file.write("mukey; bedrock_min_depth \n")
21
22  # Define counties of interest
23  counties = ['105','195','169','167','191','173','197','021','155','071','121',\
24      '185','045','161','023','005','017','091','163','015','165','139','171',\
25      '187','069','043','107']
26
27  # Loop through counties
28  for county in counties:
29      print('Working on county_' + county)
31      # Connect to database
32      db_file = r'''H:\\Research_Thesis\\THESIS\\Co_Databases\\soildb_US_\
33          2003_''' + county + r'''.mdb'''
34      user = ''
35      password = ''
36      conn_str = 'DRIVER={Microsoft Access Driver (*.mdb)};DBQ=%s;UID=%s;PWD=%s' %
37                    (db_file, user, password)
38      conn = pyodbc.connect(conn_str)
39      cursor = conn.cursor()
40
41      # Create Component_Table1
42      Make_Table1 = """ (SELECT mapunit.mukey, muaggatt.brockdepmin
43      INTO Comp_Table1 FROM (mapunit INNER JOIN muaggatt ON (mapunit.mukey \
44          = muaggatt.mukey))) """
45      cursor.execute(Make_Table1)
46
47      # delete NULL values for brockmindep
48      cursor.execute("DELETE from Comp_Table1 WHERE brockdepmin IS null")
49      print('done removing Null bedrock depth min values')
50
51      # write out all rows remaining into Python_BedrockDepth_Output.txt
52      for row in cursor.execute("SELECT * from Comp_Table1"):
53          file.write(str(row[0]) + ";" + str(row[1]) + "\n")
54
55  # Close file containing output
56  file.close()
57
58  print('Done with creating table')
```

Figure 14: *Python Code Producing Shapefile Defining Percent of Sinkholes per Risk Zone*

```python
1  # -*- coding: utf-8 -*-
2  # --------------------------------------------------------------------------
3  # 2014_8_14_Count_Analyses_NonSimplified_RasterToPolygon.py
4  # Created on: 2014-08-14 11:40:19.00000
5  #   (generated by ArcGIS/ModelBuilder)
6  # Description:
7  #   Takes a range of weighted combinations and performs a series of ArcMap
8  #   functions on the data to create a final version of clipped risk zones to
9  #   existing mapped sinkhole shapefile to determine percent of total areas of
10 #   sinkholes per risk zone
11 # Author:
12 #   Alexandra L. Todd
13 # --------------------------------------------------------------------------
14
15 # Import arcpy module
16 import arcpy
17
18 # Check out any necessary licenses
19 arcpy.CheckOutExtension("spatial")
20
21 # Overwrite Outputs
22 arcpy.env.overwriteOutput = True
23
24 # Loop through combinations
25 for i in range(2,30):
26     print "Working on Combination " + str(i)
27     import time
28     ts = time.time()
29     import datetime
30     st = datetime.datetime.fromtimestamp(ts).strftime('%X-%m-%d %H:%M:%S')
31     print st
33     # Define local variables:
34     Combo = "H:\\Research_Thesis\\THESIS\\SSURGO_EditedData\\Combo_Maps\\Combo"
35         + str(i) + "\\combo" + str(i) + "_8_6"
36     Sinkholes = "H:\\Research_Thesis\\THESIS\\SSURGO_EditedData\\Sinkholes\
37         _Hubbard_StatePlaneNVA_meters_DissolvedIntoOne.shp"
38     Combo_Reclassified = "H:\\Research_Thesis\\THESIS\\SSURGO_EditedData\\Combo\
39         _Maps\\Combo" + str(i) + "\\combo" + str(i) + "_rc"
40     Combo_Polygon = "H:\\Research_Thesis\\THESIS\\SSURGO_EditedData\\Combo_\
41         Maps\\Combo" + str(i) + "\\combo" + str(i) + "_polygon.shp"
42     Combo_Sinkholes = "H:\\Research_Thesis\\THESIS\\SSURGO_EditedData\\Combo_\
43         Maps\\Combo" + str(i) + "\\combo" + str(i) + "_sinkholes.shp"
44     Combo_Sinkholes_Dissolved = "H:\\Research_Thesis\\THESIS\\SSURGO_Edited\
45         Data\\Combo_Maps\\Combo" + str(i) + "\\combo" + str(i) + "_sinkholes\
46         _dissolved.shp"
47
48     print "Reclassifying Combination " + str(i)
49     # Process: Reclassify into integer risk categories
50     arcpy.gp.Reclassify_sa(Combo, "Value", "0 1 1;1 2 2;2 3 3;3 4 4;4 5 5;5 \
51         6 6;6 7 7;7 8 8;8 9 9;9 10 10;10 11 11;11 12 12;12 13 13;13 14 \
52         14;14 15 15", Combo_Reclassified, "DATA")
53
54     print "Converting Raster to Polygon for Combination " + str(i)
55     # Process: Convert raster to polygon
56     arcpy.RasterToPolygon_conversion(Combo_Reclassified, Combo_Polygon, \
57         "NO_SIMPLIFY", "VALUE")
58
```

```python
59      print "Clipping Combination " + str(i)
60      # Process: Clip polygon to existing mapped sinkhole region
61      arcpy.Clip_analysis(Combo_Polygon, Sinkholes, Combo_Sinkholes, "")
62
63      print "Dissolving Combination " + str(i)
64      # Process: Dissolve based on risk category
65      arcpy.Dissolve_management(Combo_Sinkholes, Combo_Sinkholes_Dissolved, \
66          "GRIDCODE", "", "MULTI_PART", "DISSOLVE_LINES")
67
68      print "Done With Combination " + str(i)
```