Identifying Karst Sinkholes from LiDAR-Derived Topographic Depressions[1]

Karst sinkholes are a major natural hazard, and approximately 20 percent of the USA is prone to sinkhole development. Information on existing sinkholes is critical in evaluating sinkhole hazards and understanding mechanisms leading to formation of sinkholes, especially catastrophic cover-collapse sinkholes. LiDAR provides accurate and high-resolution topographic information and has been used to improve the delineation of sinkholes in many karst regions. However, LiDAR data generates large quantity of topographic depressions; identifying sinkholes from these depressions through manual visual inspection can be slow and laborious, necessitating the need for automated approaches.

This dataset contains morphological characteristics of LiDAR-derived topographic depressions from three counties (Bourbon, Woodford, and Jessamine) in the Bluegrass Region of Kentucky. This dataset consists of 22,884 records with 10 variables for each record, split 80/20 for training and testing.

The variables describe three dimensional characteristics of the depressions. The surficial shape (i.e., a polygon representing the closed contour of the depression on the surface) of a depression was characterized by three variables, **shape length (perimeter)**, **shape area** , and **compactness**. The compactness, *C*, measures how close a shape resembles a circle. One way to measure the compactness is (Cole, 1964)

[1] Thanks to Junfeng Zhu, Adam Nolte, Nathan Jacobs, and Ming Ye for providing the data for this assignment.

$$C = \frac{A}{A_c}, \tag{1}$$

where A is the area of the polygon and A_c is the area of the smallest circle circumscribing the polygon. Values of the compactness ranges from 0 to 1 and equal 1 when the polygon is a perfect circle. Natural sinkholes tend to have a circular shape with high compactness value.

To capture characteristics of the depressions in the vertical direction, the statistics of depths in each depression and statistics of slopes of a 9 m (30 ft) ring (ring slope) surrounding each depression were calculated. The variables selected for characterizing depths were **maximum depth**, **mean depth**, **depth standard deviation**, and **total depth**. When multiplied by the cell size of the elevation raster, the total depth variable becomes the volume of the depression. The variables selected to describe the ring slope were **mean slope** and **slope standard deviation**. Lastly, the relationship between the depth and the surface area of a sinkhole is computed as the **depth index** (D_i), which is defined as

$$D_i = D_{max}/\sqrt{A/\pi} \tag{2}$$

where D_{max} is the maximum depth and A is the surficial area (i.e., the area of the polygon on the surface). The depth index reflects the slope of the depression by assuming a sinkhole as an upside down cone (Miao et al., 2013).