The 2022 World Congress on
Advances in Civil, Environmental, & Materials Research (ACEM22)
16-19, August, 2022, GECE, Seoul, Korea

Prediction of karst subsidence and sinkhole collapse using logistic regression and random forest methods

*Yong Je Kim[1)] and Boo Hyun Nam[2)]

[1)] *Department of Civil and Environmental Engineering, Lamar University, Beaumont, TX 77710, USA*
[1)] ykim3@lamar.edu
[2)] *Department of Civil Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do, 17104, Korea*
[2)] boohyun.nam@khu.ac.kr

ABSTRACT

Karst geohazards such as sinkholes and subsidence create substantial structural problems in buildings and surface infrastructure, resulting in significant economic losses. Sinkhole occurrence in east-central Florida (ECF) is becoming more frequent and hazardous due to urban growth that increases groundwater pumping. Predicting sinkhole/subsidence prone areas is an imperative task in engineering risk management and land-use planning. The purpose of this study is to construct means of sinkhole susceptibility assessment using two widely applied methods, namely logistic regression (LR) and random forest (RF) classification, and then to create the sinkhole susceptibility map of ECF. First, the subsidence and sinkhole locations are prepared based on the subsidence incident report database of the Florida Geological Survey (FGS). Then, the maps of twelve (12) contributing factors are prepared in raster format with ArcGIS Pro, and the relationship between subsidence and sinkhole locations and the factors is considered for susceptibility mapping analysis. Sinkhole factors used in this study include hydraulic head difference between surficial and confined aquifers, groundwater recharge rate, soil permeability, overburden thickness, surficial aquifer system (SAS) thickness, intermediate aquifer system (IAS) thickness, proximity to karst features, geology, land use/land cover, lithology, geomorphology, and average annual precipitation. Subsequently, the database is randomly divided into two subsets: 70% for training and 30% for testing. Using the training and testing sets, two susceptibility models using LR and RF methods are constructed, assessed, and compared. The prediction capability of the two susceptibility models is assessed and compared using the receiver operating characteristics (ROC) curve.

[1)] Assistant Professor
[2)] Associate Professor