The study of the vulnerability to high-impact weather using web-crawling and big-data analysis in Korea

Jaein Jung (1), Seungbum Kim (2), and Kyungjun Lee (3)

(1) Korea, Republic Of (jungji@korea.kr), (2) Korea, Republic Of (kmaskim@korea.kr), (3) Korea, Republic Of (leekj15@korea.kr)

The weather is closely affecting our lives. To analyze the effect of various weather phenomena on our lives, news articles containing certain keywords, such as heavy rain, heavy snow, heat waves, fog, strong winds, severe cold weather, are collected from the Korean portal site Naver (https://www.naver.com). Using the KoNLPy package in Python, we tokenize collected sentences into phrases and count the frequencies of weather-related words. We find that the extent of damage being collected varies depending on the combination of search keywords and the results are widely gathered when searching for meteorological phenomena with the word that means weak effect such as 'inconvenience' or 'loss'. As a result, the detailed effects of weather events on agriculture, livestock and health are listed.