

## Prediction of Heat Wave over South Korea using Deep Learning with Teleconnection Factors

Yeonsu Lee (1), Cheolhee Yoo (1), Dongjin Cho (1)

<sup>1</sup> Ulsan National Institute of Science and Technology, 50 UNIST-gil, Ulsan, 44919, Korea  
Email: [leeysu0423@unist.ac.kr](mailto:leeysu0423@unist.ac.kr); [yoclhe@unist.ac.kr](mailto:yoclhe@unist.ac.kr);  
[djcho@unist.ac.kr](mailto:djcho@unist.ac.kr)

**Abstract:** Because of global warming, the frequency of heatwave over the world has been increasing. According to the National Oceanic and Atmospheric Administration (NOAA) 2018 report, the recent 5 years are the hottest years among 1880-2018 records. The heatwave has severe effects on human mortality and our society. Thus, the accurate and early prediction of the heatwave is needed for response. South Korea is located in East Asia and a peninsula surrounded by water. The summer weather of South Korea is hot and humid which is affected by East Asian Summer Monsoon (EASM). In previous research, the various teleconnection patterns are picked for drivers of a heatwave over the Korea Peninsula. For instance, the tropic sea surface temperature (SST) over the El Niño regions, the western North Pacific and eastern North Pacific, show high correlations with the maximum temperature over Korea in the summer season. The anomaly of tropic SST is well known as a trigger of the annual variation of El Niño-Southern Oscillation (ENSO). Also, the relation of winter North Atlantic Oscillation (NAO) and spring snow cover, snowmelt over Eurasia can lead to the extremely hot summer of South Korea. Our purpose is to investigate the teleconnection factors with the heatwave and predict the frequency of heat waver over South Korea using Convolutional Neural Network (CNN). For statistical forecasting, we need data for a long period. Thus, we used the snowmelt data from the Global Land Data Assimilation System (GLDAS) Noah Land Surface Model which are in 1.0-degree resolution and range from January 1979 to the present. GLDAS uses data assimilation techniques for integrating satellite- and ground-based observational products. Also, we used the European Space Agency (ESA) Soil Moisture Climate Change Initiative (CCI), the data collected by six active and passive sensors on the satellites from 1979 to 2010, as observation data to validate the GLDAS reanalysis data. The SST data from Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) are used. Such input data have different domains; snowmelt and soil moisture over Eurasia, SST over the western and eastern North Pacific and North Atlantic. Those also have different period; SST over North Atlantic for winter, and the others for spring. Daily surface maximum temperature (Tmax) from July to August over South Korea measured by 59 Automated Synoptic Observing System (ASOS) stations from Korea Meteorological Administration (KMA) are used to calculate the total number of heatwave events as a target variable. CNN is applied to make a seasonal forecasting model for heatwave and Leave-one-out Cross-Validation (LOOCV) is used to evaluate the model. Besides, the results of Random Forest and Rotation Forest are used for comparison of model performance. This study shows the possibility of predicting the intensity of heatwave over South Korea with preceding teleconnection factors over the world.

**Keywords:** Heat wave, Teleconnection, Convolutional Neural Network