Flood risk assessment model under climate change impacts using Machine Learning algorithms

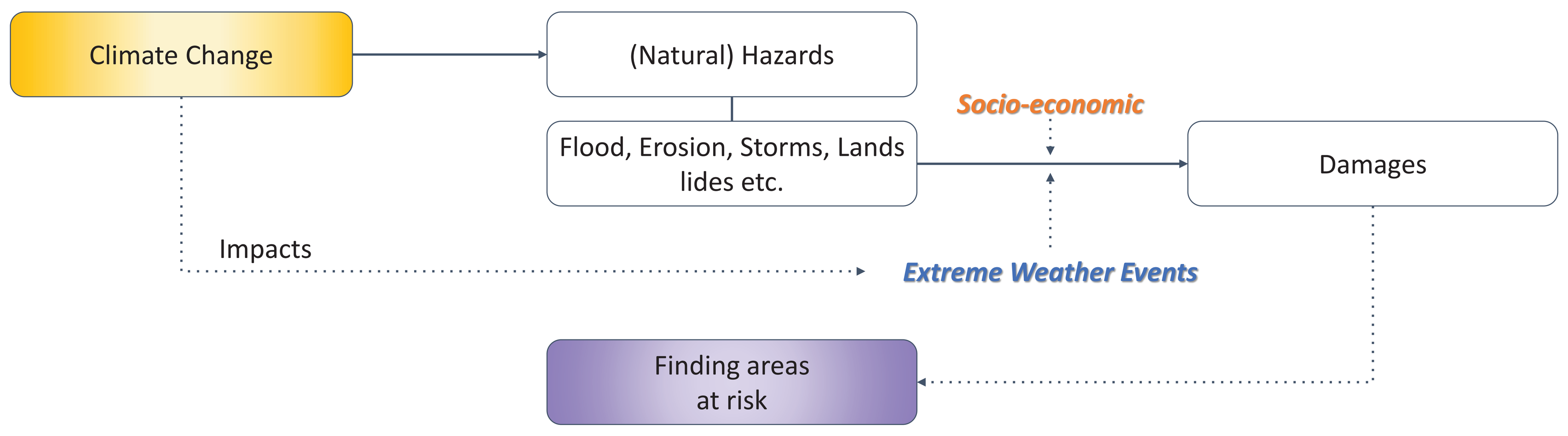
Sang-Jin Park[1], Dong-Kun Lee[1,2]

[1]Interdisciplinary Program in Landscape Architecture & Integrated Major in Smart City Global Convergence, Seoul National University, Rep. of Korea.

[2]Department of Landscape Architecture and Rural System Engineering, Seoul National University, Rep. of Korea.

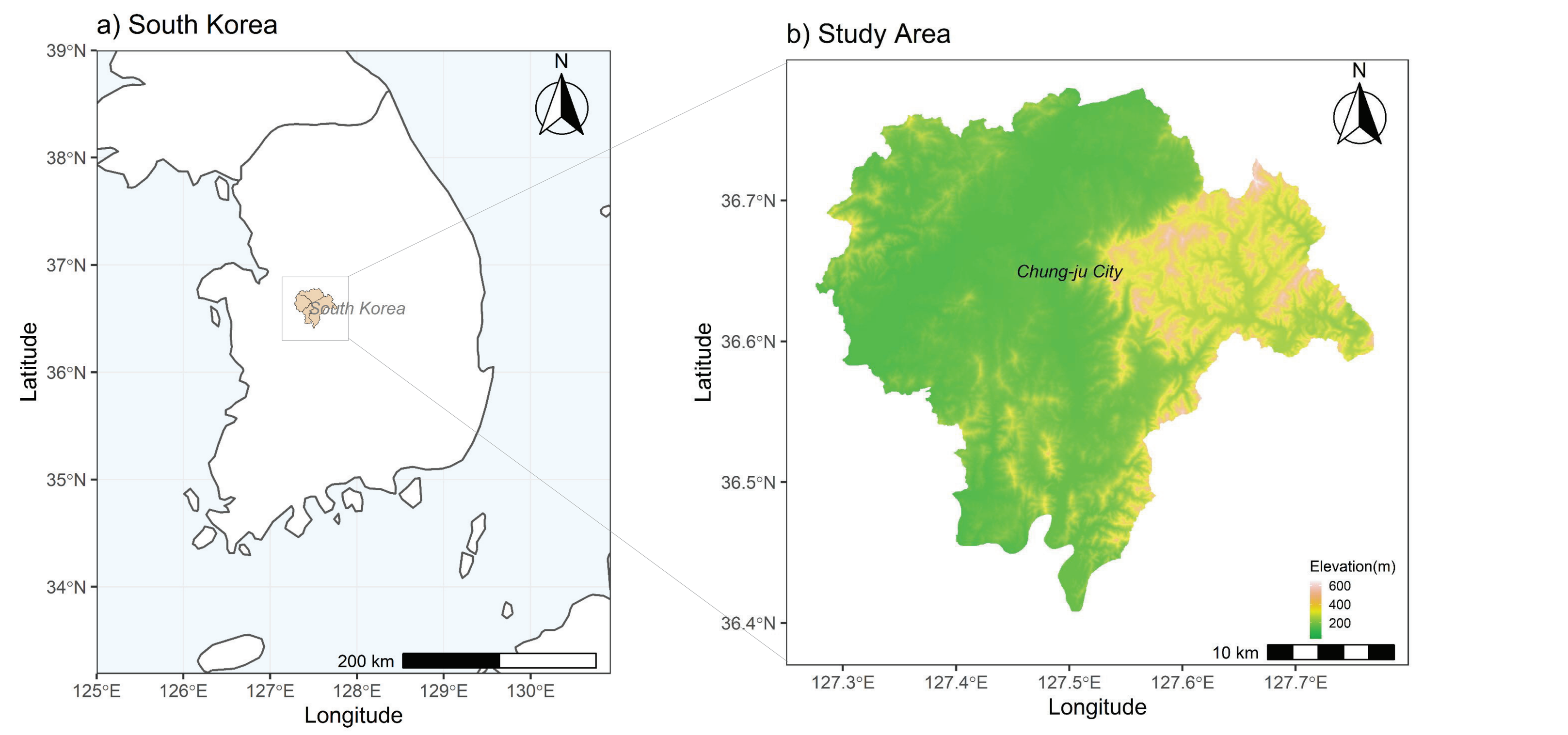
Sang Jin Park
PhD Student, Interdisciplinary Program in Landscape Architecture
E-mail : parksang8705@gmail.com

1. Background and Goal



Recently, natural hazards have been more unpredictable with increasing frequency and strength due to climate change. South Korea is no exception. Various natural disasters, such as heatwaves, torrential rains, and typhoons, have become frequent, and the direct and indirect losses caused by these disasters have been increasing significantly every year. In the future, with rapid global warming, the abnormal climate is expected to become worse. Consequently, natural disasters will become more catastrophic, inflicting critical damage on human society. Thus, a series of hazard prevention plans are necessary, and the first step would be to find risk areas. The objective of this study is to find vulnerable areas that could be damaged by flooding caused by typhoons and heavy rainfall.

2. Study Scope

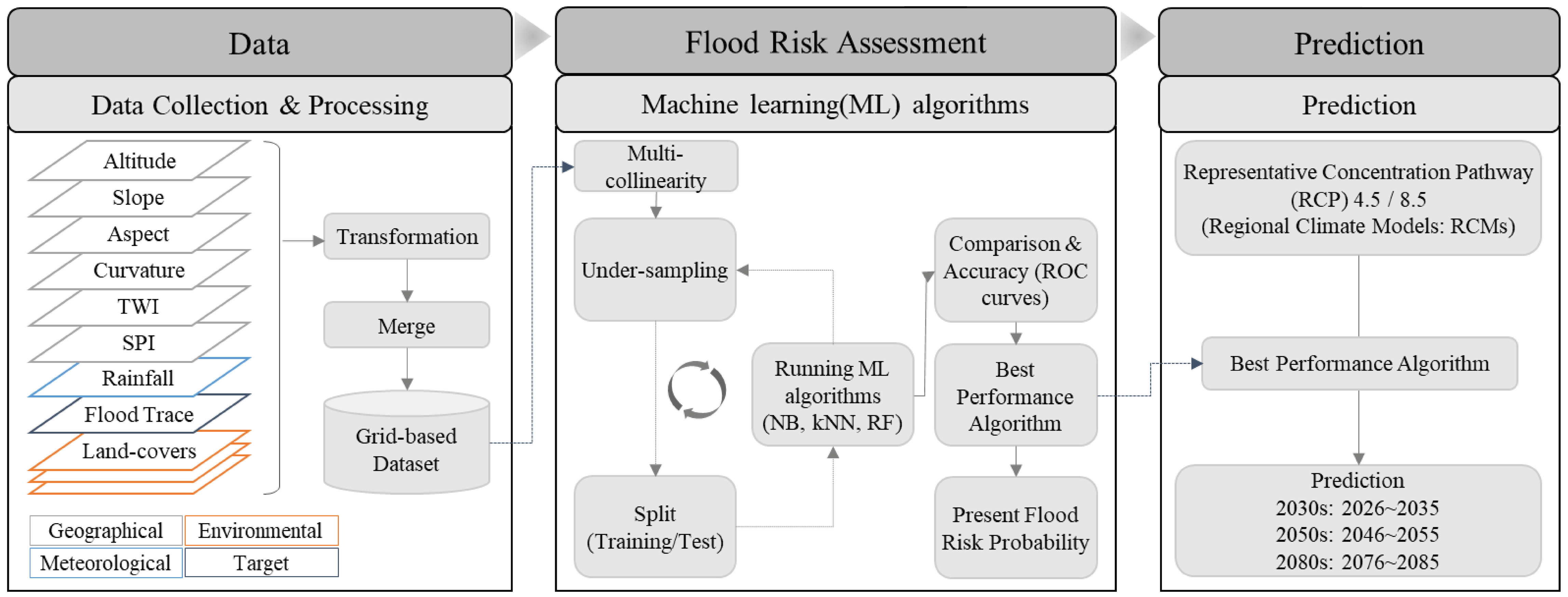


We aimed to select a region with the largest frequency of floods during a certain period within a certain range to develop a basic model. Therefore, a region in Cheongju, Chungcheongbuk Province, where frequent floods occurred in many areas within a certain range due to heavy rains with a daily maximum of 281 mm from July 16 to 17, 2017, was selected as the study area.

3. Data

Data		Source	Type	Period
Altitude		Ministry of Environment	Grid	2019
Slope		Calculated by using Altitude	Grid	2019
Aspect		Calculated by using Altitude	Grid	2019
Curvature		Calculated by using Altitude	Grid	2019
Topographic Wetness Index		Calculated by using Altitude	Grid	2019
Stream Power Index		Calculated by using Altitude	Grid	2019
Distance to river			Grid	2019
Land covers	Impervious area ratio	Ministry of Environment (https://eng.me.go.kr/eng/web/main.do)	Polygon	2018
	Agriculture area ratio			
	Forest area ratio			
Rainfall		Korea Meteorological Administration (https://web.kma.go.kr/eng/index.jsp)	ASCII	2017.07.16 ~17
Flood Trace		Korea Land and Geospatial Informatrix Corporation (https://www.lx.or.kr/eng.do)	Polygon	2017.07.16 ~17

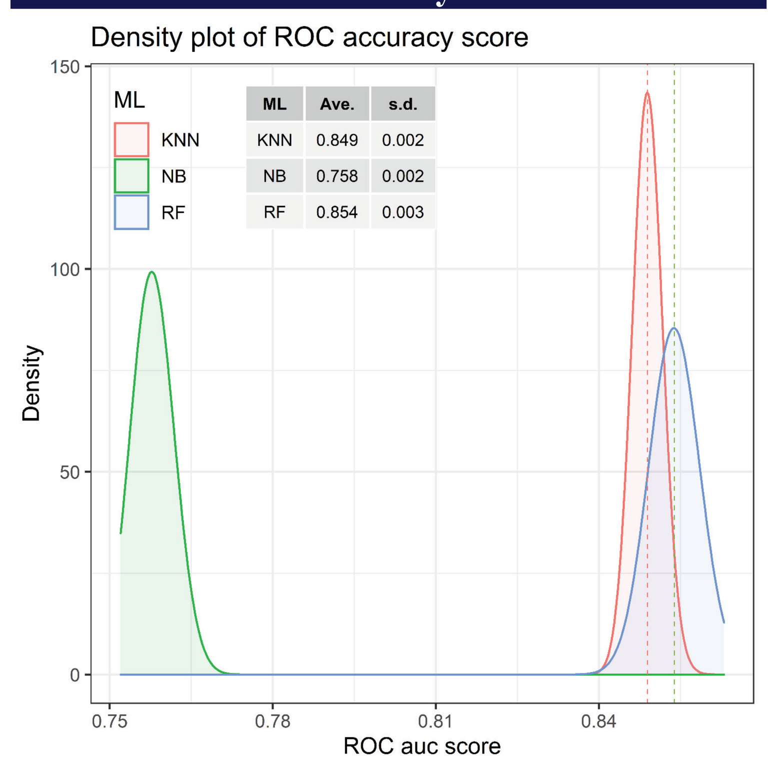
4. Method and Model Process



The flow of research is divided into three main stages. 1) Data collection and processing steps required for flood risk assessment. 2) Using the processed data set, adjusting variables in consideration of multi-collinearity, and calculating flood risk. The flood risk probability was performed using three different machine learning techniques which area Naive-bayes(NB), k-Nearst Neighbor(kNN), and Random Forest(RF), and the model with the highest performance was selected to calculate the flood risk probability. 3) To predict future flood risk, changes in climate factors. For changes in climate factors, the RCP scenario was used, and in detail, the precipitation data from five different regional climate models (RCMs) were used, and the future potential flood risk was predicted.

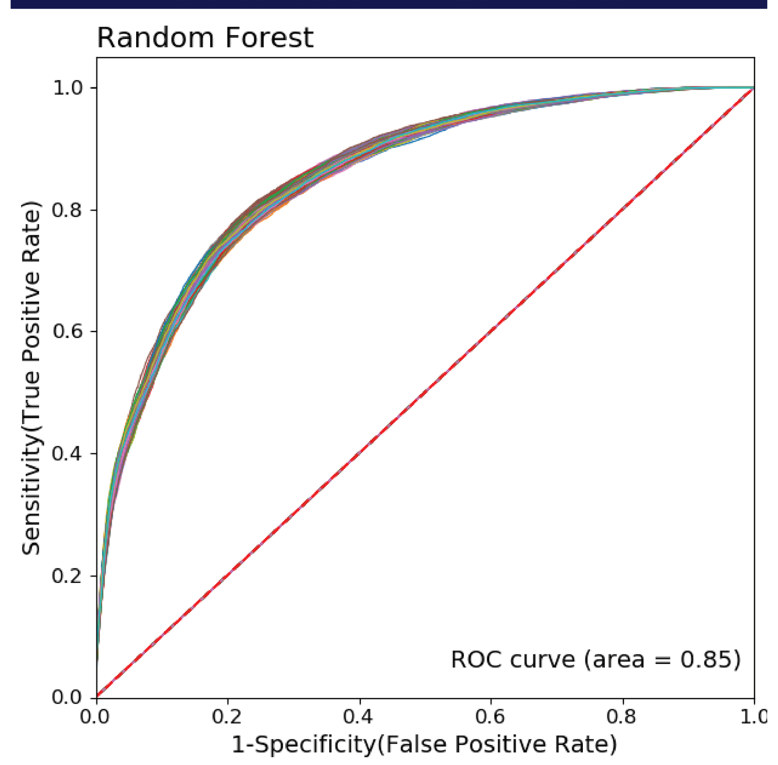
5. Results

ROC accuracy scores



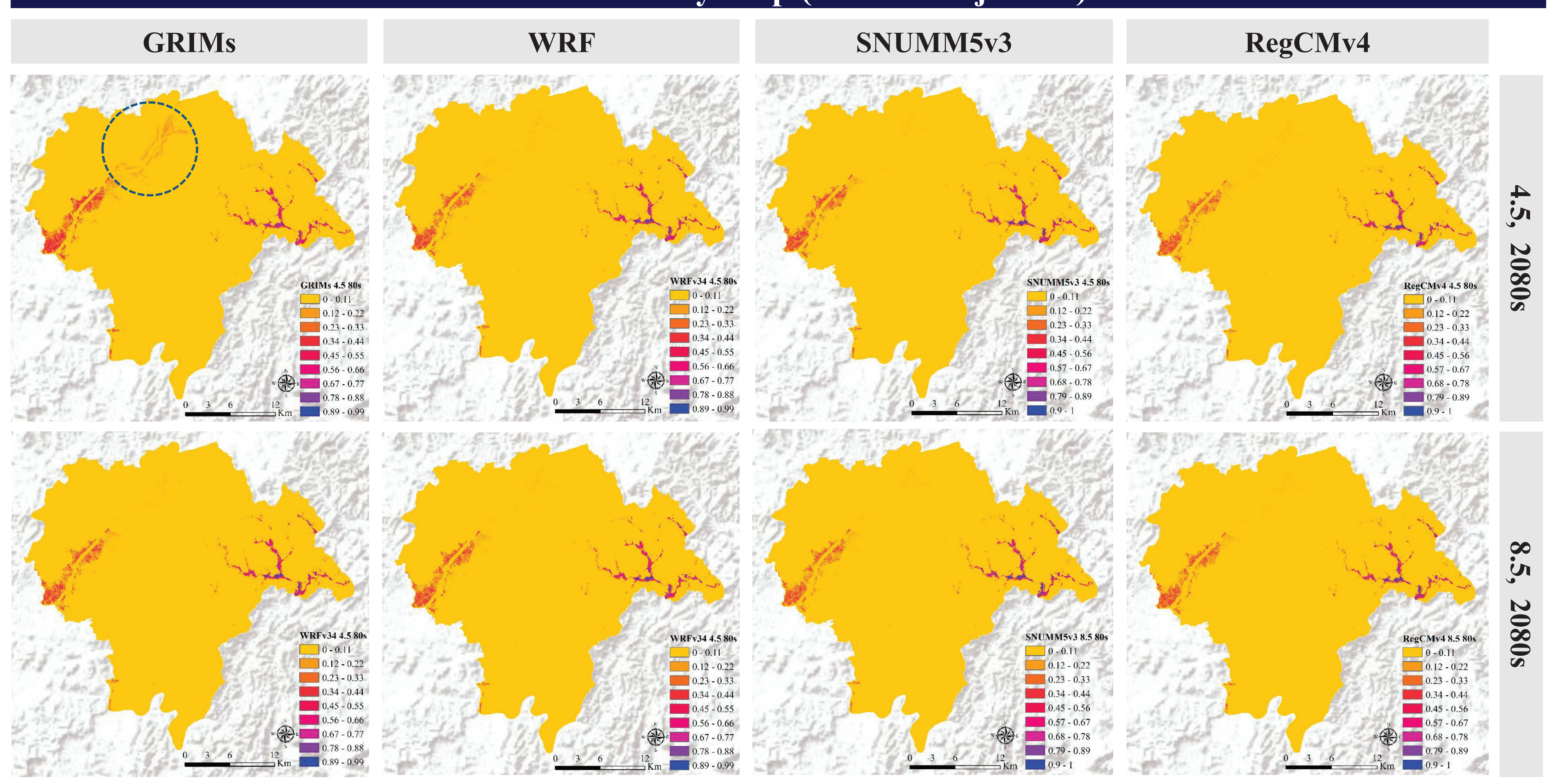
As a result of running the three ML algorithms for the flood risk analysis, the accuracy of the resulting average ROC curves of each algorithm were NB (0.758), kNN (0.849), and RF (0.854).

ROC curves of Random Forest

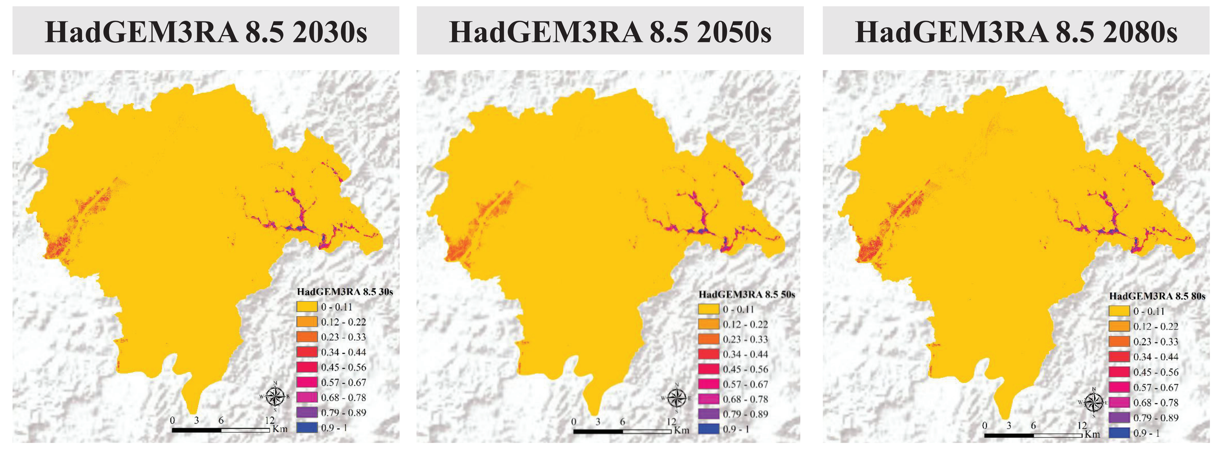


By using the RF, the ROC curves plot was made as shown the Figure upper side, which means that better performance as the shape of the graph curves toward the upper left corner.

Risk Probability Map (Future Projection)

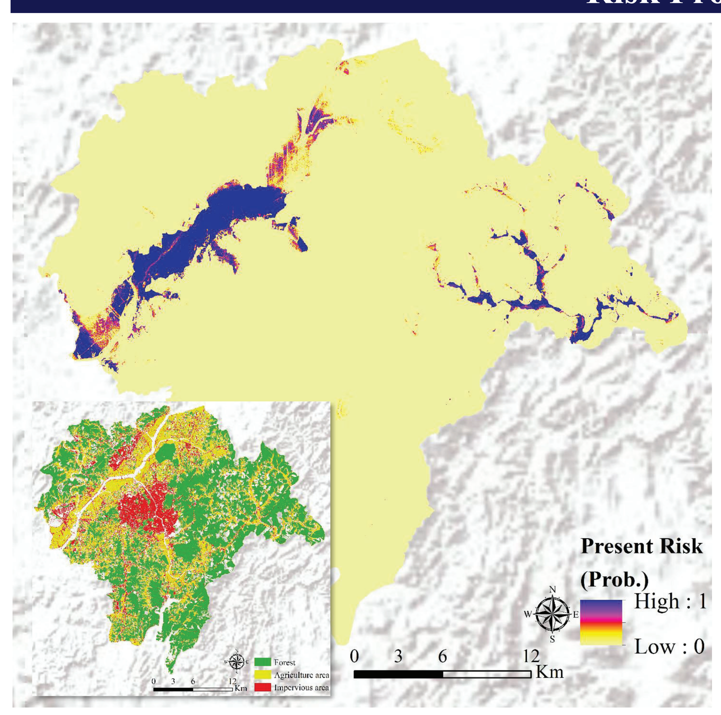


Risk changes in the future depending on the regional climate model (RCMs: WRF, SNU-RCM, RegCM4, CCLM) and target years (2080s: 2076~2085) in the RCP 4.5/8.5 scenarios. The figure shows the future potential risk changes in the 2030s, 2050s, and 2080s for RCP 4.5/8.5, according to the four RCMs. In general, risk increases from the RCP 4.5 to the RCP 8.5 among the four RCMs, and the risk of flooding increases very slightly, imperceptibly to the naked eye. Also, there are no big differences among the 8 results from RCMs. In particular, as a result of GRIMs 4.5 2080s, more several areas were estimated to occur flooding compared to the RCMs. Although there were no big differences among the results, it is a critical point that the model indicated the same areas will generally be more vulnerable.



In general, the risk of flooding increases very slightly, imperceptibly to the naked eye, from the 2030s to the 2080s in the HadGEM-3RA regional climate scenario.

Risk Probability (Present)



A risk probability map with a gradational color distribution was developed, as shown the Figure left side, based on the results of the model. The higher probabilities indicate areas at a higher risk of compounding events such as heavy rainfall. Compared to the DEM (Digital Elevation Model) of this area, the probability of flooding was high in the low-lying area, and in fact, a lot of flooding occurred in that area. Also, the most of high risk to flooding areas was estimated to occur near the agricultural areas.

※ This work was supported by Korea Environment Industry & Technology Institute(KEITI) through The Decision Support System Development Project for Environmental Impact Assessment, funded by Korea Ministry of Environment(MOE)(No. 2020002990009)