Ph.D. Dissertation of Engineering

Assessment of disaster risks induced by climate change, using machine learning techniques

머신러닝 기법을 활용한 기후변화 영향에 따른 재해 리스크 평가

August 2022

Graduate School of Seoul National University
Interdisciplinary Program in Landscape Architecture
Integrated Major in Smart City Global Convergence Program

Sang Jin Park

Assessment of disaster risks induced by climate change, using machine learning techniques

Advisor: Dong Kun Lee

A dissertation submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Interdisciplinary Program in Landscape Architecture and Integrated Major in Smart City Global Convergence Program in Seoul National University

August 2022

Sang Jin Park

Approved by Thesis Committee

Chair	_____(Seal)
Vice Chair	_____(Seal)
Examiner	_____(Seal)
Examiner	_____(Seal)
Examiner	_____(Seal)

Thesis and Dissertation Deposit Agreement

Under this Agreement, I represent and warrant that my dissertation of thesis (the "Work") does not infringe the intellectual property rights, including copyright, of any third party. I grant the Seoul National University (the "SNU") certain rights as follows.

Title	Assessment of disaster risks induced by climate change, using machine learning techniques
Degree	Mater ☐ / Ph.D. ▣
Major	Interdisciplinary Program in Landscape Architecture and Integrated Major in Smart City Global Convergence Program
Student ID No.	
Tel./Mobile No.	

1. I hereby consent to authorize the SNU to reproduce, distribute, display and transmit the Work over the internet through the SNU library.
2. I hereby grant to SNU the royalty-free right to use for online service.
3. I agree that SNU may, without changing the content, translate the Work to any medium or format.
4. If I grant to others copyright ownership, I will notify a relevant office in SNU so as to make the Work available to the public or not.
5. Under Article 46 of the Korean "Copyright Act", SNU will exploit the Work when the holder of author's property right grants authorization, and will not be obligated to take legal action related to any intellectual property rights in the Work.

Date : **2022 . 8 . 8 .**

Author : Sang Jin Park

Seoul National University

Abstract

Assessment of disaster risks induced by climate change, using machine learning techniques

Sang Jin Park

Interdisciplinary Program in Landscape Architecture and Integrated Major in Smart City Global Convergence Program in Seoul National University

Graduate School of Seoul National University

Supervised by Professor Dong Kun Lee

Climate change is an urgent threat to our generation. Natural hazards have become more unpredictable, occurring more frequently and with greater force, due to climate change. Natural disasters in Korea are mostly caused by meteorological events. The total damage caused by disasters in the last ten years is attributed mainly to typhoons (49%) and heavy rain (40%). Therefore, risk management, which analyzes and evaluates hazard risk related to heavy rainfall such as flooding and landslides, is needed to prepare for the long term. Also, effective monitoring and detection responses to climate change are critical for predicting and managing threats to hazard risks.

Therefore, the main research questions of this thesis are as follows: 1) How to predict future potential risks in a complex situation due to climate change considering various factors, 2) And what kind of efforts are made to reduce such risks? Is it sustainable? First of all, to assess the future risk of multiple hazards such as coastal flooding, landslide, 1) this study analyzed the present risk by using multiple machine learning (ML) algorithms that have been

widely used in recent studies as part of probabilistic approaches, and future risks were estimated by considering the forecasted rainfall according to different representative concentration pathway (RCP) climate change scenarios and regional climate models. Secondly, to evaluate the effectiveness of adaptation strategies to respond to disaster risks posed by climate change impacts, 2) this research analyzed the effectiveness and sustainability of structural measures such as green space and seawall, which are widely used and play an important role as countermeasures against coastal flooding, by dividing into several adaptation pathways.

The results of this study identify future at-risk areas and can support decision-making for risk management and can guide disaster reduction and management measures, including land use planning and decision-making processes.

Keyword: Climate change impacts, Disaster Risk Reduction (DRR), RCP scenario, sea level change, Integrated Coastal Zone Management (ICZM), Landslide susceptibility, Nature based Solution (NBS), adaptation strategy

Student Number: 2019-39641

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 2

Chapter 3

Chapter 4

Publications

Portions of this dissertation have been reprinted with permission, from previously published materials. Chapters 2 and 3 have been reprinted from *Environmental Research Letters*, *Geomatics, Natural Hazards and Risk*, respectively. Chapter 4 was submitted at *The Science of Total Environment*.

1. Park, S. J. and Lee, D. K. (2020). Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. Environmental Research Letters. 15(9).

2. Park, S. J. and Lee, D. K. (2021). Predicting susceptibility to landslides under climate change impacts in metropolitan areas of South Korea using machine learning. Geomatics, Natural Hazards and Risk, 12(1), 2462–2476.

3. Park, S. J., Sohn, W. M., Piao, Y., and Lee, D. K. (2022). Adaptation strategies to future coastal flooding: performance evaluation of green and grey infrastructure in South Korea. The Science of Total Environment (submitted version).

Acknowledgments

1. This work was supported by the Korea Environmental Industry and Technology Institute (KEITI) grant funded by the Korea government (ME) (No. 2014001310007)

2-3. This work was supported by the Korea Environment Industry & Technology Institute (KEITI) through the Decision Support System Development Project for Environmental Impact Assessment, funded by the Korea Ministry of Environment (MOE) (No. 2020002990009).

Chapter 1. Introduction

1. Background

Climate change is an urgent threat to our generation. Natural hazards have become more unpredictable, occurring more frequently and with greater force, due to climate change (Berz et al., 2001; Kundzewicz et al., 2014; UNISDR, 2015). Urban areas are where population, human activity, and artificial structures are spatially concentrated. In the event of a disaster, human and economic damage tends to increase compared to other areas (Huppert and Sparks, 2006). On the other hand, natural disasters in South Korea are mostly caused by meteorological events (Yoon et al., 2016; Azam et al., 2017; Han et al., 2018). The total damage caused by natural disasters in the last ten years is attributed mainly to typhoons (49%) and heavy rain (40%) (Ministry of the Interior and Safety, 2016). Flood damage in South Korea will increase further due to climate change in the future. Therefore, risk management, which analyzes and evaluates hazard risk related to heavy rainfall such as flooding and landslides, is needed to prepare for the long term.

Coastal areas are threatened by hazards such as flooding, erosion, and storms (Klein et al., 1999; Nicholls and Cazenave, 2010; Saxena et al., 2013; Lilai et al., 2016) and will be more vulnerable in the future because of climate change impacts such as sea level rise and extreme weather events (Klein et al., 1999; Lilai et al., 2016; Vousdoukas et al., 2018). Furthermore, the number of people living in coastal areas globally is expected to increase from 1.8 to 5.2 billion by the 2080s (IPCC, 2014; Bhable, 2015; Neumann, 2015). As coastal systems continue to become more socially and

environmentally complex, the cost of damage from coastal hazards due to climate change impacts will also rise (Kleint et al., 2001; IPCC, 2007; Szlafsztein & Sterr, 2007; Balica et al., 2012). In particular, South Korea is a peninsula with several large cities situated along the coast and 27.5% of its total population living in coastal areas (Oh et al., 2020). Thus, a series of hazard prevention plans and the identification of risk areas are crucial for these coastal areas (Tran et al., 2008; Kourgialas & Karatzas, 2011).

However, existing research on coastal flooding has primarily used quantitative indices to characterize risks. These index studies have focused primarily on analyzing vulnerabilities that may indicate the relative risks encountered by coastal areas, but do not specifically calculate the actual risks. These previous studies also only analyzed current vulnerability or risk and did not analyze future risks however, several others have analyzed future risks. Although they analyzed the risk of coastal hazards using statistical and physically based methods, they did not consider uncertainty by comparing multiple models to predict future hazards.

It is also very important what measures should be taken to reduce risk and damage in response to coastal flooding. Measures for best coastal management that consider both cost and efficiency to optimize limited resources will become essential in the future (Ferreira et al., 2019). These efforts can mitigate risks in a proactive manner before disasters occur (FEMA, 2018; Reguero et al., 2020). As a result of reviewing related previous studies on the effectiveness of measures and strategies to reduce the risk of coastal inundation, these prior studies lacked consideration of spatiotemporal analysis. The impacts and vulnerabilities of climate change may differ depending on the spatial and temporal characteristics of countries

under present and future climatic conditions (IPCC, 2007; 2014; 2022). Therefore, it is necessary to consider analysis through spatiotemporal down-scaling in order to compare and evaluate the effectiveness of adaptation strategies to mitigate the risk of coastal inundation.

Along with coastal flooding, landslides are a major cause of serious damage to life and property worldwide (Malamud et al., 2004; Gomez & Kavzoglu, 2005; Lee & Pradhan, 2007; Garcia-Rodriguez et al., 2008; Yilmaz, 2010; Pham et al., 2020). In the future, when the impacts of climate change become more severe, sudden heavy rains could cause more damage due to landslides and flooding (Yılmaz, 2009). Also, studies of the vulnerability or susceptibility to landslides have been conducted worldwide. Previous research has mostly focused on identifying the factors There have been many great previous studies, the contents of the studies were similar only differed in the target sites and methodologies. Most of them focused on evaluating current susceptibility or sensitivity based on the conditions of the target site, collecting data on those factors, and analyzing vulnerability or sensitivity through statistical models as shown in Table 1 in Chapter3. Few studies have focused on predicting the future risk of landslides. In the field of disaster research, predictions or future prospects are important in terms of disaster management (UNISDR, 2015).

2. Purpose

Therefore, the main research questions of this thesis are as follows: 1) How the future risks of multiple disasters due to climate change would be changed considering various factors, 2) And what kind of efforts are effective to reduce such risks? To answer the

above two questions, this study set the scope of the study to include heavy rains as the cause of the greatest damage in South Korea and coastal flooding and landslides as disasters caused by heavy rains.

As a method to answer the first question, a widely used machine learning algorithm was used as part of a probabilistic approach in a recent study to evaluate the potential risk of future disasters such as coastal flooding and landslides. In addition, when analyzing potential future risks, it was estimated to consideration with the predicted rainfall according to various representative concentration pathway (RCP) climate change scenarios and sea level rise forecasts obtained through spatio-temporal detailed modeling. Next, the effectiveness and sustainability of structural measures such as green spaces and seawalls (Dong et al., 2020; Jeong et al., 2021), which are widely used and play an important role as countermeasures against coastal inundation, are compared and analyzed by dividing into several adaptation pathways.

The results of this study are thought to be able to support decision-making on Korea's Integrated Coastal Zone Management (ICZM) by identifying potential hazardous areas in the future and proposing sustainable adaptation strategies. Furthermore, it is expected to contribute to disaster reduction and management measures, including land use planning and decision-making processes (Akgun, 2012; Nsengiyumva et al., 2018; Dou et al., 2019b, 2020b).

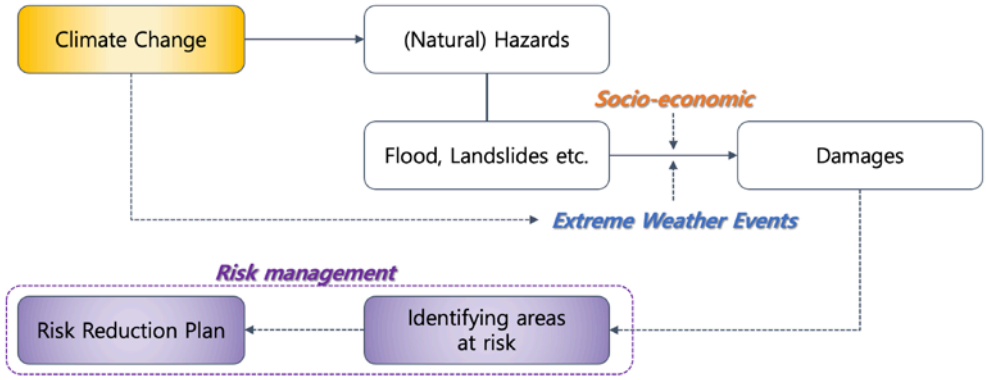

Figure 1. Conceptual framework of the research

Chapter 2. Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms

1. Introduction

Existing research on coastal flooding has primarily used quantitative indices to characterize risks. The index method expresses the vulnerability of an area using arithmetic operations that incorporate classified factors affecting hazards. Several studies have generated a coastal/composite vulnerability index for calculating and assessing of coastal hazards in different study areas by using different variables (Dwarakish et al 2009, Sankari et al 2015, Pantusa et al 2018, Sahana and Sajjad 2019). These index studies have focused primarily on analyzing vulnerabilities that may indicate the relative risks encountered by coastal areas, but do not specifically calculate the actual risks.

These previous studies also only analyzed current vulnerability or risk and did not analyze future risks; however, several others have analyzed future risks. They calculated the future risk of coastal hazards by estimating water level heights (Wahl et al 2016, Vousdoukas et al 2016). Unlike previous reports, this study quantitatively calculated coastal hazard risks and predicted future risks by considering the occurrence of compounding probabilistic events such as extreme precipitation and the rising tidal ratio. Although they analyzed the risk of coastal hazards using statistical and physically-based methods, there are some factors that could not be addressed. These studies did not obtain a spatial distribution of risk (Wahl et al 2016) and analyzed continental scales that are difficult to apply to regional scales (Vousdoukas et al 2016). Also,

they did not consider uncertainty by comparing multiple models to predict future hazards. In addition, rainfall is a very important factor in flooding but they herein focus only on changes in water level.

In short, this study analyzed the actual risk probability, not a relative vulnerability, using a coastal flooding risk analysis that considers rainfall events as well as tidal levels, because the risk to coastal areas of heavy rainfall depends on the tide (Van Den Hurk et al 2015, Eilander et al 2020). Additionally, future coastal flooding risks were estimated by considering the actual rising rate of the tide and the forecasted rainfall according to different representative concentration pathway (RCP) climate change scenarios and regional climate models. Multiple machine learning (ML) algorithms that have been widely used in recent studies as part of probabilistic approaches were used to probabilistically calculate the coastal flood risk. The results of this study identify future at-risk areas and can support decision-making for integrated coastal zone management (ICZM) in South Korea by identifying which areas require hazard prevention plans.

2. Materials and Method

2.1 Study Area

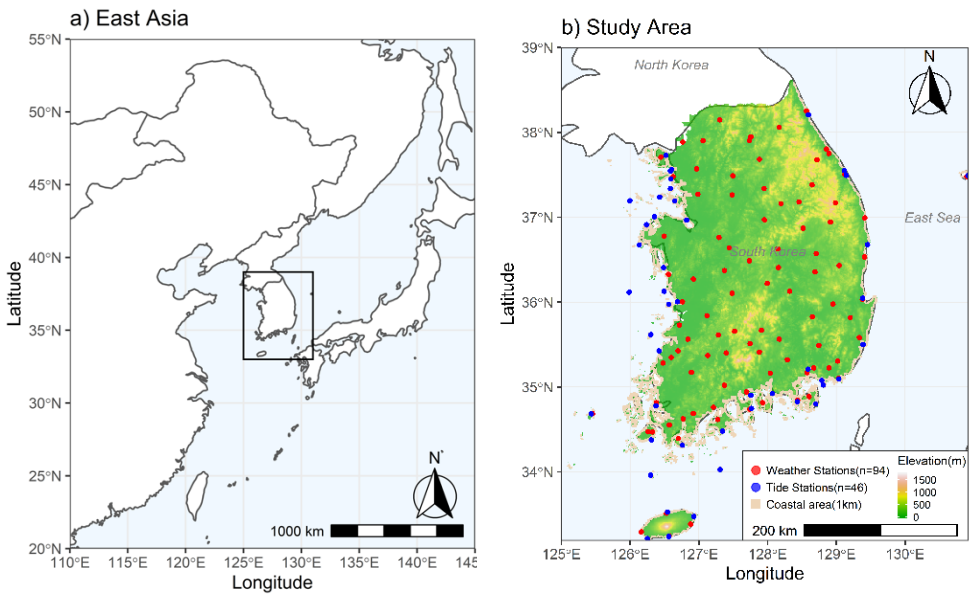

Figure 1. (a) Countries in East Asia, (b) South Korea and the study area (Coastal area: 1 km, Geodetic Datum: WGS84)

In this study, the spatial coverage is South Korea (33–38◦ N, 125–131◦ E). Summer in South Korea is generally hot and wet and typhoons that occur frequently in July and August bring heavy rainfall to coastal areas. The heaviest rainfall in this region was recorded in mid–July and mid–August when the daily average rainfall ranged between 220 mm and 322 mm. The maximum daily rainfall recorded for the period of 1973–2010 was 870.5 mm in Gangneung, South Korea, on August 31, 2002 (Korea Meteorological Administration 2011).

The total length of the coastline of Korea is 14 962.8 km and the spatial scope of this research was set up 1 km from the coastline, following the ‘coastal management law’ in Korea. The coastline along the East Sea is monotonous, and the water depth is generally deep. The West Sea has a shallow coastline with an average depth of 44 m.

The coastline of the South Sea is complex and contains numerous islands and harbors. Figure 4 shows the locations of the 68 weather observatories and 46 tide observatories in Korea.

2.2 Machine learning algorithms

A coastal hazard risk analysis was implemented using three ML algorithms: k-nearest neighbor (kNN), random forest (RF), and support vector machine (SVM). Previous studies have frequently compared these three machine learning techniques in their prediction methods (Harefa and Pratiwi 2016, Potdar and Kinnerkar 2016; Lopez-Serrano et al 2016, Danades et al 2017, Thanh Noi and Kappas 2017). The results of these algorithms were subsequently compared.

The kNN, proposed by Cover and Hart (1967), is an easy-to-implement supervised ML algorithm that is as simple as the Naive-Bayes Classifier (Jadhav and Channe 2016). The proximity of data points to one another affects the results of the algorithm (Bhavsar and Ganatra 2012, Kim et al 2012). In this study, the analysis was performed by setting k as 5, 10, and 15. The analysis was the best when k was set to 5. The RF algorithm is an ensemble learning method operated by constructing multiple decision trees during a training period, and it is frequently used in research along with ML techniques such as SVM and neural networks. More details related to RF are described by Breiman (2001). Setting the number of trees and depths is important; herein, these values were set to 10 and 3, respectively. The SVM algorithm, proposed by Cortes and Vapnik (1995), is a versatile ML algorithm in that it can be classified in unlabeled datasets. When dividing two sets to create the classifier

hyperplane (Rebentrost et al 2014), various classifiers such as linear and non-linear forms can be generated according to the characteristics of the data. Therefore, this algorithm is a high-performing technique used to analyze real-world data (Tong and Koller 2001). Set- ting the kernel function is important in SVM since it helps to overcome shortcomings related to linear separability (L´opez-Serrano et al 2016). In this study, we used the radial basis function (RBF) kernel function, which makes non-linear classifiers

These ML algorithms were used in this study to compensate for the shortfalls of each individual algorithm (Hao et al 2019). Additionally, by using ML algorithms, we could consider complex and diverse influencing factors caused by climate change. The results of the algorithms were compared through approximately 700 iterations to reduce the uncertainty of the model itself.

2.3 Method

2.3.1 Data

The variables used in the analysis (tide, rainfall, elevation, slope, urban area, and grassland) were selected based on previous literature reviews (Mahendra et al 2010; Sankari et al 2015, Ashraful Islam et al 2016, Giardino et al 2018, Pantusa et al 2018), while rainfall, which was unaddressed in previous studies, was used to consider compounding events in this study. All data obtained for the risk analysis were transformed into a 1 km^2 grid because the raw data consisted of different points and polygons. All of the data were obtained from Official Korean Government websites (table 1). They were organized into a data table by day and grid. In addition, a map showing data for coastal flooding traces was obtained and used as the

labeling data for classifying a machine learning algorithm. The grids on the map where coastal flooding occurred were labeled '1' (173 cases), or '0'(224,053 cases) if no flooding occurred.

Table 1. List of variables

Data (abbreviation)	Source (abbreviation, Website)	Type	Period
Mean tide(T)	Korea Hydrographic and Oceanographic Agency (KHOA, www.khoa.go.kr/eng/)	Point	2002–2014
Daily maximum rainfall (R)	Korea Meteorological Administration (KMA, https://web.kma.go.kr/eng/index.jsp)		
Elevation (E)	Ministry of Environment (ME, https://eng.me.go.kr/eng/web/main.do)	Grid	2010
Slope (S)			
Urban area (U)			
Grassland (G)			
Coastal flood trace (CF)	Korea Land and Geospatial Informatrix Corporation (LX, www.lx.or.kr/eng.do)	Polygon	2002–2014
RCP precipitation data	KMA Climate Information Portal (KMA, http://climate.go.kr)	Ascii	2002–2018

2.3.2 Coastal risk analysis

The entire data table, in which all variables were organized by day and grid, was under-sampled for the risk analysis due to imbalanced data that included more cases of non-flooding than flooding. By under-sampling the data, the risk analysis could be performed (He and Garcia 2009). The under-sampled data were subsequently split into training (70%) and test (30%) datasets, after which the risk analysis was implemented using kNN, RF, and SVM.

The procedure from the under-sampling to running the algorithms was repeated 1200 times. After running the three ML algorithms, the results were compared using the receiver operating characteristic (ROC) accuracy scores and curves. ROC curves are mainly used to assess model accuracy, and the model is judged by the relationship between the false positive rate (1-Specificity) and the true positive rate. The risk probability was calculated, and risk maps were constructed using the results obtained from the three ML algorithms with the highest accuracy scores.

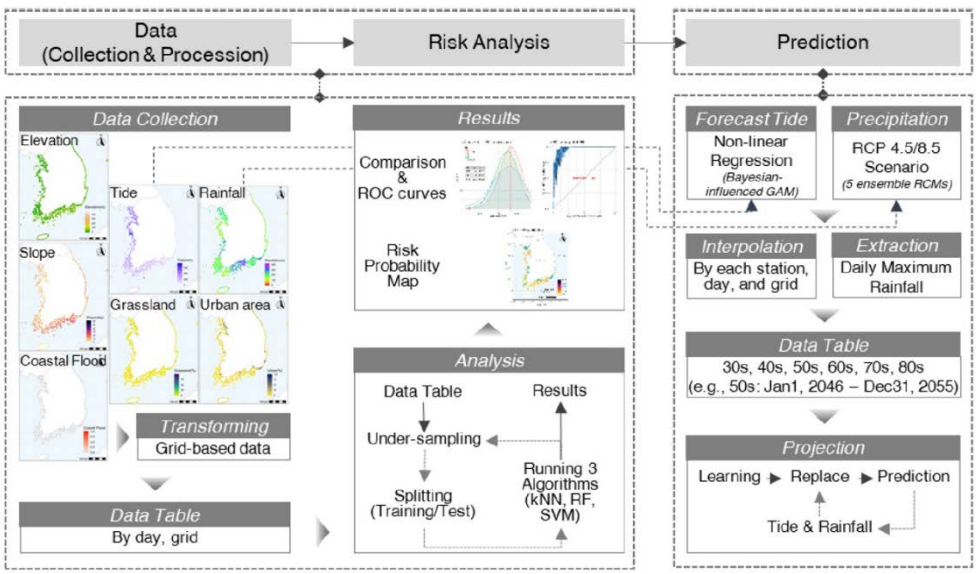

Figure 2. The sequence of the method from data collection to prediction.

2.3.3 Prediction

The future risk probability was predicted using the highest performance algorithm. To predict future risks under the impacts of climate change, the continuous variables (rainfall and tidal level) were forecasted daily into the future. These were used to evaluate future coastal flooding risks

Initially, the rainfall data used was the RCP AR5 scenario (4.5, 8.5) precipitation data produced from five different regional climate

models (RCMs): CCLM, HadGEM3-RA, RegCM4, SNU- RCM, and WRF. These models were produced by the Regional Climate Detailing Project in East Asia (CORDEX-EA: Coordinated Regional Downscaling Experiment—East Asia, source: http://cordex-ea.climate.go.kr/cordex). These were obtained from the KMA climate information Portal (table 1), and KMA uses the HadGEM3-RA as its principal data. Daily maximum rainfall amounts for the different RCP scenarios of the RCMs were used, and monthly average rainfall values were calculated. Next, the tide was forecasted using real tidal range data obtained from a number of tidal observation stations. A Bayesian-influenced generalized additive model (GAM) was used to accurately forecast the tidal data as a sine-shaped curve with repetitive rising and falling trends. The Bayesian-influenced GAM, based on Bayes' theorem, performs regression while keeping functions smooth as a non-linear regression (Wood, 2020). The past tidal pattern for each tidal station was analyzed and future tidal values were calculated and organized by day. The monthly average tidal values were then calculated from the daily tidal values.

Lastly, the calculated future tidal and rainfall data were then used to predict future risk probabilities. Future target periods ranged from the 2030s to the 2080s, with one division spanning 5 years (e.g. 2050: January 1, 2046–December 31, 2055). The process of running the model for prediction consisted of three steps (figure 2). (1) The kNN classifier was created similar to the manner in which it was created in the coastal risk analysis, (2) tidal and rainfall variables were each replaced by the predicted values, and (3) the replaced data table was used for prediction. This routine was repeated almost 1200 times to reduce uncertainty.

3. Results

3.1 Comparison of ML algorithms

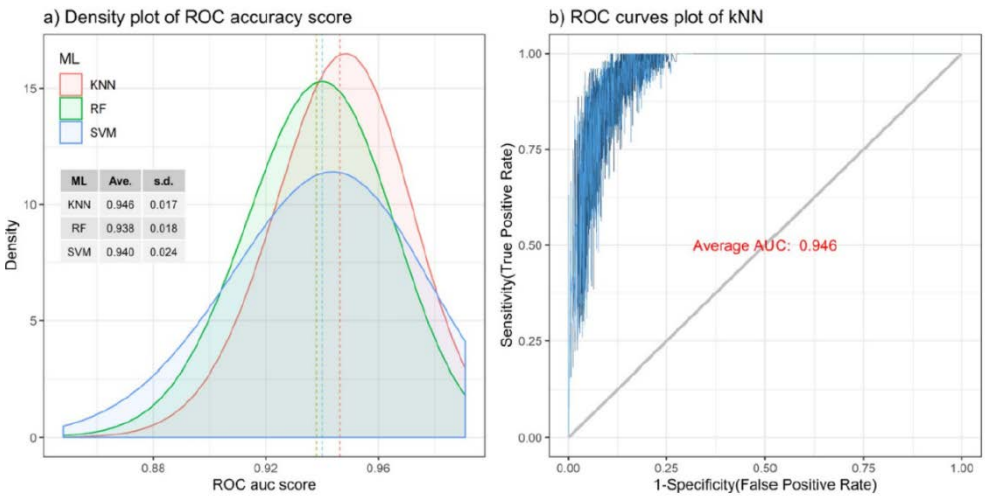

Figure 3. (a) ROC curves of accuracy scores: comparison of results of 3 ML algorithms, and (b) plot of k-nearest neighbor ROC curves.

As a result of running the three ML algorithms for the coastal risk analysis, the accuracy of the resulting average ROC curves of each algorithm were kNN (0.946), RF (0.938), and SVM (0.940), as shown in figure 3(a). The ROC curves of the other models did not appear to be low either; however, the kNN model produced relatively better results. Also, figure 6(a) shows that the kurtosis at the density of the kNN ROC accuracy is slightly higher than the others, which means that the accuracy of the kNN is not as biased compared to the others. Therefore, the kNN was used for the final risk probability mapping and future prediction analysis. Figure 6(b) shows the trade-off between the false positive rate (1-Specificity) and the true positive rate. When the curves are closer to the upper left corner, its classifier exhibits good performance.

3.2 Risk probability map

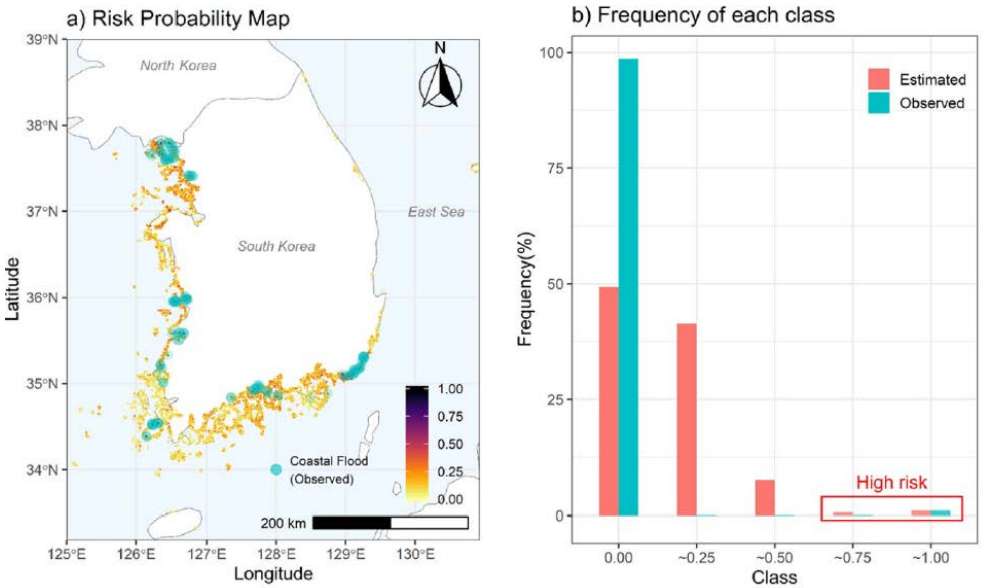

Figure 4. (a) Risk probability map (blue dots are the observed coastal flooding points, and the gradational color distribution indicates the estimated result, which is the risk probability, from the kNN algorithm), (b) frequency of each class (a comparison between the observed coastal flooding events and the estimated result of the kNN when the result is classified in five classes)

A risk probability map with a gradational color distribution was developed, as shown in figure 7(a), based on the results of the model. The higher probabilities indicate areas at a higher risk of compounding events such as high tides and heavy rainfall. The blue dot on the graph indicates where coastal flooding occurred from 2002–2014. Comparing where the actual flooding occurred and was estimated to occur, the risk probability was relatively high in the area where the actual flooding occurred. Figure 7(b) compares the frequency in percentage by each class between the actual coastal flooding point mentioned above and the risk probability calculated by the kNN. According to this result, the value calculated by the kNN model overestimated the risk class below 0.75; however, it was approximately 64.35% accurate in estimating risk probabilities above

0.5. Even if the model for calculating risk probabilities tends to overestimate the risk, areas where risk probabilities of 0.5 or higher have been derived could still be at risk in the near future.

3.3 Future risk under climate change impacts

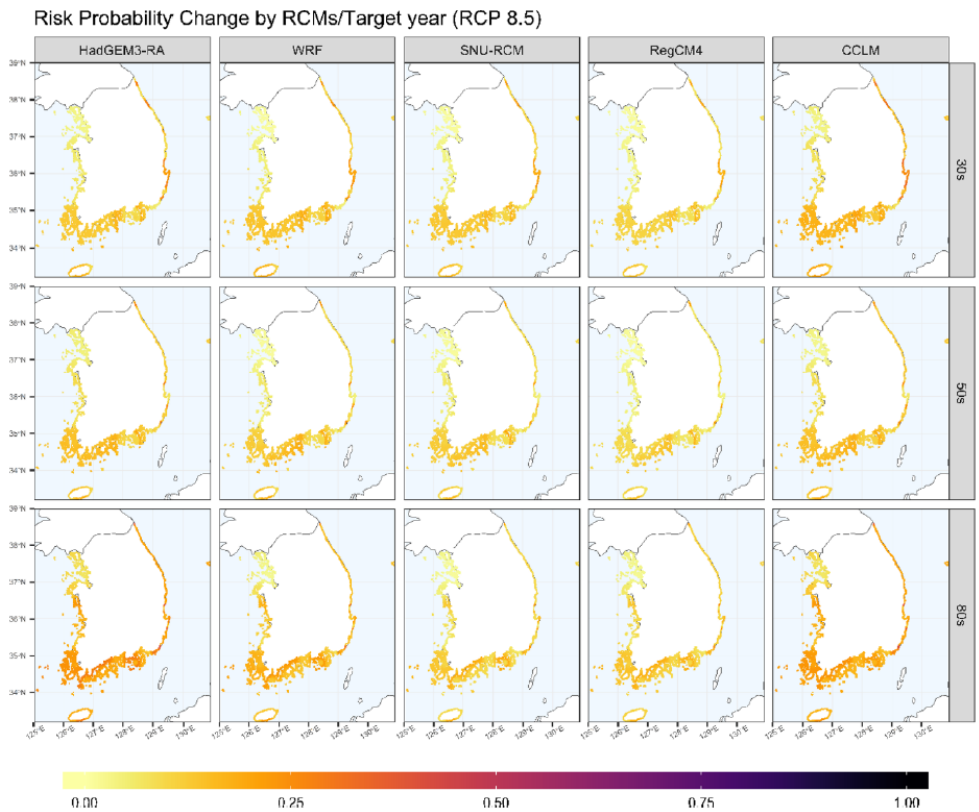

Figure 5. Risk probability changes in the future depending on regional climate model (RCMs: HadGEM3-RA, WRF, SNU-RCM, RegCM4, CCLM) and target years in the RCP 8.5 scenario.

The prediction was implemented using monthly average rainfall and tidal values as described. In the process, rainfall predicted to occur in the future was used as a density function to consider the uncertainty of future rainfall. Tidal data was input into the kNN classifier by month, the rainfall value was estimated by substituting the kernel density in one month, and the model was used to calculate the monthly predicted risk. Then, according to the comparison of the

monthly risk probability data, the risk probability increased for the months of June, July, and August for most of the RCP 4.5 and 8.5 scenarios. Based on these data, the average risk probability for each scenario (RCP 4.5/8.5; the 2030s to the 2080s) in June, July, and August was calculated to create a maximum risk probability map for the scenario. Figure 8 shows the future risk probability changes in the 2030s, 2050s, and 2080s for RCP 8.5, according to the five RCMs. In general, risk increases from the 2030s to the 2080s among the five RCMs, and in particular, from the 2050s to the 2080s for CCLM and HadGEM3-RA. Although there were differences among the results obtained using the five RCMs, the southern coastal areas will generally be more vulnerable than the eastern and western areas. The reason for these regional differences should be determined, as this could be significant for coastal zone management.

4. Discussion

4.1 Regional differences

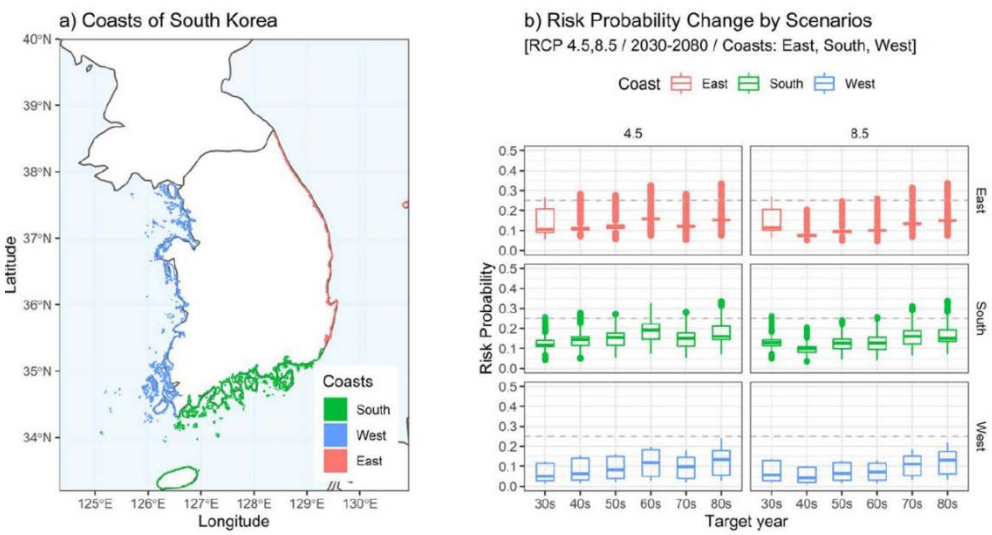

Figure 6. Comparison of the future risks faced by the three coasts: risk probability change according to the RCP scenarios (4.5/8.5) and target year (2030s–2080s) by using the result of the average regional climate models (RCMs).

The description of the impacts under different climate change scenarios in the Intergovernmental Panel on Climate Change (IPCC) report states that, as climate change progresses, the world will be affected differently by region (IPCC 2007, 2014). In Korea, which is surrounded by three seas, the geographical characteristics along the eastern, western, and southern coasts are different (figure 1). Therefore, the impacts of climate change are expected to differ, and therefore the risks associated with coastal inundation also differ.

Figure 6 displays the change in risk probability for the three seas around South Korea from the present to the future. In these graphs, the southern coast shows slightly more risk than the other two coasts from the 2030s to the 2080s at both RCP scenarios (4.5, 8.5). The risk level also exhibits an increase in the 2060s in the RCP 4.5 scenario, but the risk increases in the 2070s in the RCP 8.5 scenario. Although there is a difference in the time at which the risk increases, this suggests that the risk probabilities will increase in any scenario in the 2050s. Therefore, measures for long-term adaptation (30 or 40 years from now) should be prepared.

4.2 Significance factor

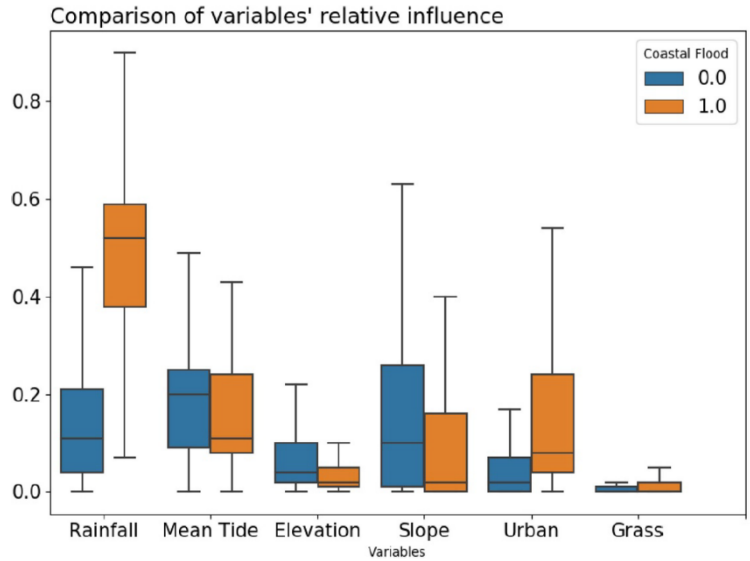

Figure 7. Comparison of the relative influence of variables according to the conditions of an occurrence and a non-occurrence of a coastal flood (0.0: not occurred, 1.0: occurred)

The average tidal values for the three bordering seas (west, south, and east) are 546.56 mm, 222.16 mm, and 23.76 mm, respectively, in the 2050s. The average elevations are 22.8 m, 50.9 m, and 43.1 m, respectively. Theoretically, the western area should be the most vulnerable since the ratio of tidal rise is higher and the average elevation is lower than the others, as shown in figure 1. However, both the southern and western areas are also at risk, though the gap will gradually increase in the future. In order to determine the reason why the risk along the southern coast was estimated to be higher than that of the other two coasts, we investigated which variable dominated the results.

We divided the original data that was used for the risk analysis into whether the coastal flood event occurred or not. Then, each variable was normalized from 0 to 1 and the results were compared. The difference in rainfall between the normalized values

according to floods that occurred was higher than the others. We inferred that rainfall is therefore a key factor compared to other variables, such as altitude or slope, in the risk analysis, as shown in figure 7 (Ward et al 2018). This also demonstrates that the water level does not fit the assumption that the western coastal area is theoretically more at risk, as shown in figures 5 and 6. Furthermore, the fact that urban areas are frequently flooded may suggest that coastal management plans such as building facilities for protection should account for the vulnerability of urban coastal areas.

4.3 Methodological implications

This study compared coastal flood risk analyses using three ML algorithms. As a result of the risk analyses, the results of the kNN analysis model exhibited the highest reliability and accuracy. This suggests a slightly different implication from that of other studies. In other studies that used various ML techniques, the accuracy of the algorithm was higher when using a ML technique such as artificial neural networks (ANN), SVM, or RF than those obtained when using kNN (Potdar and Kinnerkar 2016, L´opez− Serrano et al 2016, Thanh Noi and Kappas 2017). We concluded that the performance of the kNN in this study was slightly higher because it may have been influenced by the difference in data quality. This study used traces of actual flooding for the risk analysis, and as a result of running the model with these data, the accuracy was high. We infer that the risk analysis using kNN may be applied broadly as a quantitative technique, unlike studies with index methods, according to the spatial distribution of the flooded region data, regardless of region. In addition, the data−driven statistical method

using ML algorithms as well as kNN is useful in terms of scalability, because it can account for various influences such as compounding events and can quickly adapt to the input of new data. Therefore, this quantitative approach could be effective for risk analysis.

Moreover, we attempted to consider uncertainty by using an ensemble method such as comparing the results from the five RCMs. The climate model itself does have uncertainty, and that uncertainty increases over time (Knutti and Sedl´a˘cek 2013). Therefore, many studies consider uncertainty regarding future climate change using an ensemble approach that compares multiple models (Parker 2013). This could be the best way for decision-makers to communicate about future risks. In this study, trends were confirmed by comparing several climate models rather than one. We used a data-driven method instead of a model-driven method because it is difficult to confirm the uncertainty of future risk in a model-driven approach when using large amounts of data from the regional climate models used in the study.

5. Conclusions

Six variables were used to evaluate the future probability of coastal flooding events based on three different ML algorithms, namely kNN, RF, and SVM. All the data obtained for the model, such as tidal, rainfall, and elevation data, were converted into data in a 1 km2 grid since each raw data type consisted of different points or polygons. Using the three ML classifiers method, the risk probability was calculated and the results of the ROC curves and their accuracy scores were compared. The average accuracy score of the kNN was the highest (0.946), and a risk probability map was developed using

the results estimated by the kNN classifier. To evaluate future coastal flood risks due to climate change, tidal and rainfall data were used as a continuous value in prediction. For the RCP (4.5/8.5), daily maximum rainfall data for different RCMs from the 2030s to the 2080s (e.g. 2050s, 01.01.2046–12.31.2055) were used for model prediction and the kernel density was used as the input data for the prediction. In terms of the tidal level, the rising values of future tides were calculated by considering the rate of increase at each tidal station and forecasted using a Bayesian-influenced GAM. We estimated the future risk probability using forecasted tidal and future rainfall. As a result, the risk probability increased over time and the risk probability increased in the southern coastal areas more so than in the eastern or western coastal areas

In this study, we argue that there are significant implications. We initially found that the results of the kNN were performed slightly better than the other methods by comparing three ML algorithms. This can be attributed to the characteristics of the kNN, according to the quality of the original data. It also infers that risk analysis using a simple ML algorithm such as kNN could be applied widely, regardless of region. Next, rainfall was identified as a significant factor in this study. This means that the possibility of flooding can be increased due to the uncertainty in forecasting future rainfall patterns due to climate change. Lastly, future coastal flood risk analysis was analyzed using ensemble methods from different RCMs. We considered future uncertainty, though it might be helpful for decision-makers to communicate about future risks by providing variances and trends from different results.

As in the previous claim, the ML technique used in this analysis exerts a powerful force when reliable data is used, but the

results are not as sophisticated or deterministic as the results of a model-driven analysis, such as a hydrodynamic model. As long as there is uncertainty regarding climate change, a data-driven approach using ML may be easy for predictive analyses. Therefore, future studies could address the idea that the future tide and surge heights are calculated using a hydrodynamic model together, as in the work of Hoch et al (2019) and Muis et al (2020) to improve the quality of the results. This could also be aligned with precipitation or rainfall to consider compound flooding of discharge-tidal interactions through statistical analysis, as in the work of Eilander et al (2020). In addition, shoreline changes could be included in the analysis of coastal hazard risks, making the results more meaningful for ICZM. Furthermore, for the purposes of predicting future risk, it was generally assumed that variables other than tide and rainfall would not change over time. Geographical factors such as elevation and slope might not vary with time, but land cover such as urban areas and grasslands will change over time. Thus, according to land cover changes, the future spatial distribution of risk probability could be different. However, the rainfall was a key factor in this analysis as described previously, so the land cover change might not affect strongly the prediction. Therefore, it would be informative if a similar study could be conducted that accounts for social and economic changes in the risk analysis.

Chapter 3. Predicting susceptibility to landslides under climate change impacts in metropolitan areas of South Korea using machine learning

1. Introduction

Along with flooding, landslides are a major cause of serious damage to life and property worldwide (Malamud et al., 2004; Gomez & Kavzoglu, 2005; Lee & Pradhan, 2007; Garcia-Rodriguez et al., 2008; Yilmaz, 2010; Pham et al., 2020). In the future, when the impacts of climate change become more severe, sudden heavy rains could cause more damage due to landslides and flooding (Yılmaz, 2009). Therefore, studies related to landslide-susceptibility assessments and responses are necessary to guide disaster reduction and management measures, including land use planning and decision-making processes (Akgun, 2012; Nsengiyumva et al., 2018; Dou et al., 2019b, 2020b).

Studies of the vulnerability or susceptibility to landslides have been conducted worldwide. Previous research has mostly focused on identifying the factors that cause landslides based on the conditions of the target site, collecting data on those factors, and analyzing vulnerability or sensitivity through statistical models, as shown in Table 1. A review of previous studies indicates that, while the target sites and methodologies differed, the contents of the studies were similar (Tien Bui et al., 2012; Zare et al., 2013; Pham et al., 2016; Kumar et al., 2016; Kornejady et al., 2017; Chen et al., 2017a; Hong et al., 2017; Nsengiyumva et al., 2018; Polykretis et al., 2019; Pham & Prakash, 2019; Dou et al., 2020a; Wang et al., 2020).

Table 1. Previous landslide studies that used machine learning algorithms.

Author, Year	Study area	Method used in study
Tien Bui et al, 2012	Hoa Binh province, Vietnam	Adaptive Neuro−Fuzzy Inference System
Zare et al, 2013	Vaz Watershed, Iran	Multilayer perceptron and radial basic function
Pham et al, 2016	Uttarakhand state, India.	Naïve Bayes Trees, Support Vector Machines
Kumar et al, 2016	Indian Himalayas	Fuzzy−frequency ratio
Pham et al, 2017	Indian Himalayas	Multiple Perceptron Neural Networks
Kornejady et al, 2017	Golestan Province, Iran	Maximum Entropy
Chen et al, 2017a	Langao County, China.	Rotation forest ensembles, Naive Bayes Tree
Hong et al, 2017	Chongren area, China	frequency ratio, certainty factor, index of entropy
Nsengiyumva et al, 2018	Eastern Province, Rwanda	Spatially different criteria evaluation methods
Polykretis et al, 2019	Mediterranean catchment, Greece	Adaptive neuro−fuzzy modeling
Pham & Prakash, 2019	MuCang Chai, northern Vietnam	Bagging−based Naïve Bayes Trees
Dou et al, 2020	Northern parts of Kyushu, Japan	Support vector machine hybrid ensembles
Wang et al, 2020	Sichuan Province, China	deep belief network (DBN)

There have been many great previous studies, but most of them focused on evaluating current susceptibility or sensitivity. Few studies have focused on predicting the future risk of landslides. In the field of disaster research, predictions or future prospects are important in terms of disaster management (UNISDR, 2015). The

probabilistic statistical techniques used in previous studies (Table 1) to predict and forecast landslides in a specific vulnerable area also have the potential for future studies (Pourghasemi et al., 2020). Therefore, this study aimed to predict future landslides in areas that have not received adequate attention in previous studies by using multiple climate change scenarios. A probabilistic statistical model was constructed to estimate landslide-susceptibility because it can consider the uncertainties in the calculations used for prediction (Barría et al., 2019).

Recently, machine learning (ML) algorithms have become popular and are being used extensively for the spatial prediction of diverse types of hazards (Chen et al., 2017a). Also, data-driven models, such as machine learning models, performed better and were considered more efficient than other approaches, such as expert opinion-based methods (Goetz et al., 2015; Pham et al., 2020). In this study, landslide susceptibility was assessed using five ML algorithms widely used in previous studies: Naïve bayes classifier (NB), k-Nearest Neighbor (kNN), Decision Tree (DT), Random Forest (RF), and Support Vector Machine (SVM). Predictions of future landslides were then made by considering the probability distribution of precipitation data obtained from representative concentration pathway (RCP) climate change scenarios provided by the Intergovernmental Panel on Climate Change (IPCC) and regional climate models (RCMs) provided by the Korea Meteorological Administration (KMA). In addition, climate models have uncertainties because of uncertainty in the scenario values (Knutti & Sedláček, 2013). This is the reason many studies have used the ensemble approach to consider uncertainty by using diverse scenarios (Parker,

2013). The findings of this study can serve as a data source for formulating long-term policies for response and disaster management related to landslides.

2. Materials and Method

2.1 Study Area

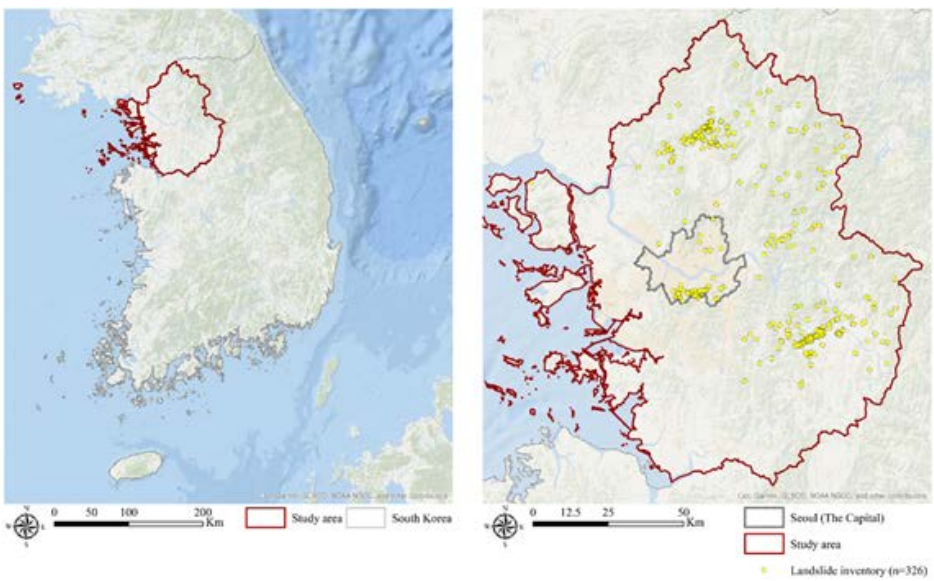

Figure 1. Study area (metropolitan area including Seoul, South Korea) and landslide occurrences (n=326, 2011–2017).

Figure 1 shows the study area and locations where landslides have occurred. This study considered the metropolitan area of South Korea (33–38° N, 125–131° E), including the capitol city of Seoul (36–38° N, 126–128° E). The metropolitan area of South Korea covers 11,851.26 km², accounting for 11.8% of the total area of Korea. At the end of October 2020, the number of residents registered in the metropolitan area was 26.3 million according to demographic data, accounting for 50.21% of the total population of Korea (Ministry of the Interior and Safety official website: https://www.mois.go.kr). The

summer climate in Korea is generally hot and humid, especially from July to August, and typhoons are common from August to September. The annual cumulative precipitation during the summer (June, July, and August) is 892.1 mm, accounting for approximately 61% of the annual precipitation. The highest recorded daily rainfall in 37 years in Seoul was 354.7 mm, which occurred on August 2, 2020 (KMA official website: https://www.weather.go.kr).

2.2 Data

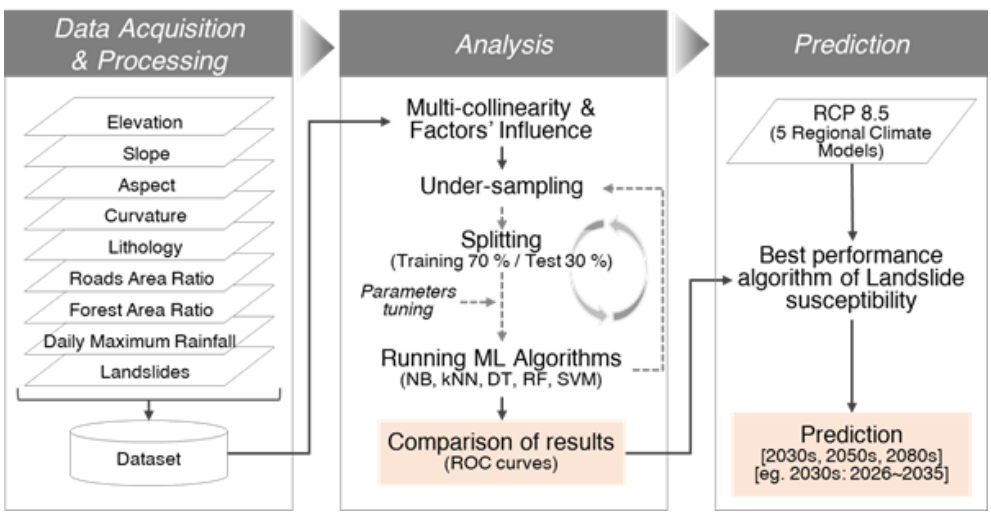

Figure 2. Research workflow

As shown in Figure 2, nine factors were used in the study. The selected factors were the same as those used for analyzing the impact or sensitivity of landslides by other studies (Tien Bui et al., 2012; Zare et al., 2013; Pham et al., 2016; Kumar et al., 2016; Kornejady et al., 2017; Nsengiyumva et al., 2018; Polykretis et al., 2019). Topographic (elevation, slope, aspect, curvature), geologic (lithology), environmental (road area ratio, forest area ratio), and meteorological (daily maximum precipitation) data pertaining to

these variables were collected for analysis from National Geographic Information Institute, Korea Institute of Geoscience and Mineral Resources, Korean Meteorological Administration, and National Disaster Management Research Institute (Table 2). By collecting data on these eight factors, as well as the landslide inventory, four topographic factors based on the digital elevation model (DEM) were created using a 10 m grid (Dou et al., 2015), and others including environmental factors were created using a 250 m grid. All data was resampled to a 250 m grid in consideration of the area of the target site. Because the dataset was created by matching the rainfall data with the landslide occurrence dates from the inventory, the size of the dataset used for analysis was the number of grids multiplied by the number of occurrences. Table 2 shows the source, type, and period of each factor. In addition, Figure 3 shows the mapping of variables used in the study, with July 27, 2011 mapped as an example for Daily Maximum Precipitation (DMP).

2.3 Landslide factors analysis

To control unnecessary factors used in the analysis prior to a landslide susceptibility assessment (LSA), multi-collinearity analysis was used to find the relationship among the factors (Bui et al., 2019, Dou et al., 2019a), and the use of information gain ratio (IGR) to determine the degree of factor influence on the results was analyzed. This is because multi-collinearity and influencing factor analysis using IGR affect the results and accuracy of the model (Zhou et al., 2018). Variance Inflation Factor (VIF) and tolerances were used to calculate the accuracy of multi-collinearity. Then, the

influence of each factor was determined using the IGR technique (Zhou et al., 2018).

Table 2. Data acquisition for the landslide-susceptibility analysis.

Category	Factor (Abbreviation)	Source	Type	Period
Topographical	Elevation (E) Slope (S) Aspect (A) Curvature (C)	*NGII	Grid	2014
Geological	Lithology (L)	**KIGAM	Polygon	-
Environmental	Roads area ratio (RA)	NGII	Polyline	2018
	Forest Area Ratio (FA)		Polygon	2017
Meteorological	Daily Maximum Precipitation (DMP)	***KMA	Points	2011-2013
	Precipitation from Regional Climate Models (*used for prediction)		Grid	2026.1.1-2085.12.31
Target	Landslides (LS)	****NDMI	Points	2011-2013

*NGII (National Geographic Information Institute)
**KIGAM (Korea Institute of Geoscience and Mineral Resources)
***KMA (Korean Meteorological Administration):
****NDMI (National Disaster Management Research Institute)

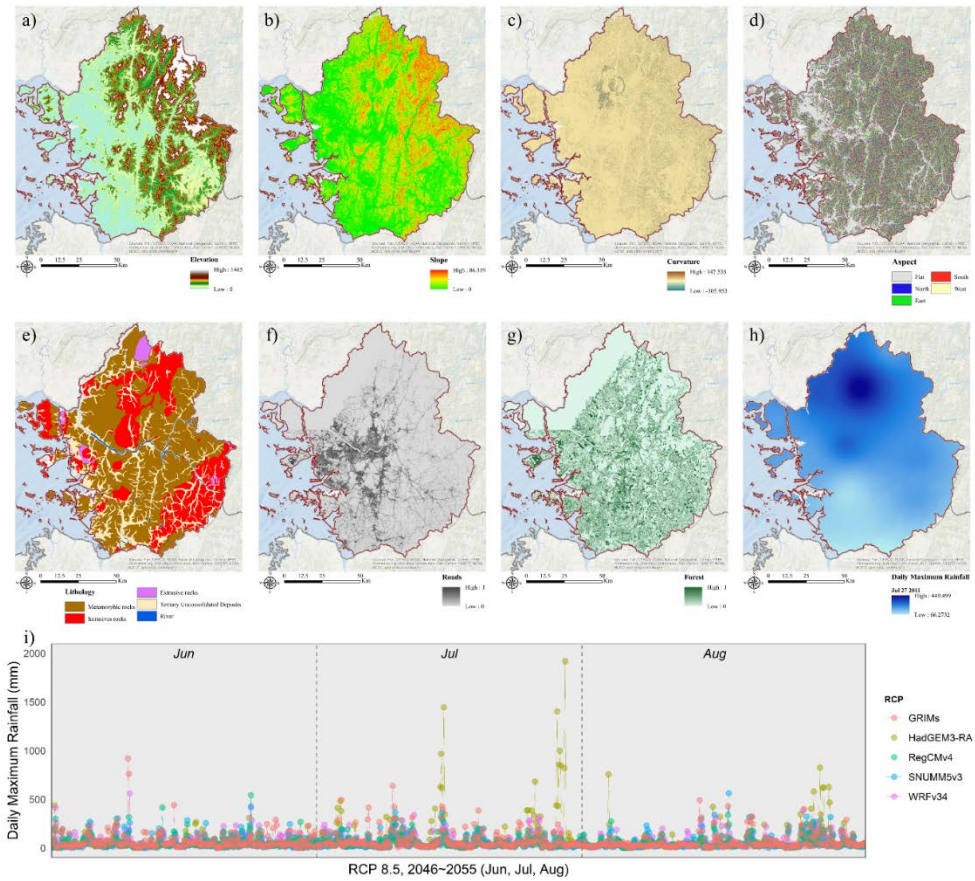

Figure 3. Maps of variables used in this study: a) Elevation, b) Slope, c) Aspect, d) Curvature, e) Lithology, f) Roads area ratio, g) Forest area ratio, h) Daily maximum rainfall (ex. July 27 2011), i) Time series graph in summer season (Jun to Aug) of daily maximum rainfall in RCP 8.5 scenario according to the five different Regional Climate Models.

2.4 Machine learning algorithms and validation

Each of the five ML algorithms used in this study had their own characteristics. NB is a stochastic-statistical method based on Bayes' rule, where prior probability is used to estimate the posterior probability. Bayes' theorem is stated mathematically as follows: $P(A|B) = P(A)P(B|A)/P(B)$. In that equation, $P(A)$ and $P(B)$ is the prior probability, $P(B|A)$ is the likelihood, and $P(A|B)$ is the posterior probability. kNN was developed by Cover and Hart (1967),

is easy to run, and is as simple as NB (Jadhav & Channe, 2016). The designation of the number k, which is the proximity of the data points to one another, is important because it affects the result of the algorithm (Bhavsar & Ganatra, 2012; Kim et al., 2012). DT, developed by Breiman et al. (1984), is a popular ML algorithm resembling a tree and based on decision tree theory. It is useful for decision-making because it provides a simple representation of the results (DeFries & Chan, 2000). RF is an algorithm that is often used in studies that use ML, combined with algorithms such as SVM and neural networks (e.g., Breiman, 2001). SVM, devised by Cortes and Vapnik (1995), is a multi-purpose algorithm that can classify unlabeled datasets. It identifies and analyzes the characteristics of data clusters.

Using the receiver operating characteristic (ROC) curve score, the results of the LSA derived from the five ML algorithms were compared. ROC analysis was mainly used to assess model performance (Pham et al., 2016), the relationship between the false positive rate (1-Specificity), and the true positive rate determining the model's performance. The closer the average AUC (area under the ROC curve) is to 1, the higher the accuracy of the model (Chen et al., 2018).

2.5 LSA using different algorithms

The landslide inventory was divided into inventories for landslide occurrence and non-occurrence. Because there was a considerable difference in frequency between the occurrence and non-occurrence areas, under-sampling was performed based on the

occurrence area (He & Garcia, 2009). The analysis was then conducted by dividing the data into a training set and a test set with a ratio of 70:30 (Dou et al., 2015, 2019b, 2020; Zhou et al., 2018). Tens of thousands of iterations were required to analyze all of the grids because under-sampling was performed prior to grid creation. Landslide susceptibility was assessed using five ML algorithms that have been used widely in recent years: NB, kNN, DT, RF, and SVM. Additionally, these five algorithms were used to account for the uncertainty of the model (Hao et al., 2019). Results were obtained through an analysis involving approximately 50,000 iterations.

2.6 Predicting landslide susceptibility

Susceptibility was predicted using the highest performing algorithm. Precipitation data from RCP climate change scenario 8.5, obtained from five different RCMs (GRIMs, HadGEM3-RA, RegCM4, SNURCM, and WRF), were used as the main variables for prediction. RCP climate change scenarios are scenarios using radiative forcing to measure the amount of carbon emissions that are the main cause of climate change, and there were four scenarios (2.6, 4.5. 6.0, 8.5). The higher the number, the higher the carbon emission scenario, with 8.5 indicating the scenario with carbon emission continuously carried out at the current level (IPCC, 2014). Precipitation data for these five RCMs were obtained from the Regional Climate Detailing Project in East Asia (CORDEX-EA: Coordinated Regional Downscaling Experiment-East Asia, http://cordexea.climate.go.kr/cordex). The climate information portal of the KMA provided the RCM data (Table 2). We used the daily maximum rainfall amounts for the different RCP

scenarios of the RCMs, and their probability distributions were used as inputs to determine the future susceptibility to landslides. The temporal targets for forecasting are the 2030s, 2050s, and 2080s, and daily maximum rainfall amounts for 10 years before and after each year were used (Figure 2).

3. Results

3.1 Multi-collinearity and influencing factor analysis

Table 3. Multi-collinearity analysis among factors.

Factors	All factors		Without 'Slope'	
	VIF	Tolerance	VIF	Tolerance
Elevation	3.657	0.273	1.872	0.534
Slope	5.013	0.199		
Aspect	3.039	0.329	2.505	0.399
Curvature	1.217	0.822	1.000	1.0
Lithology	1.697	0.589	1.696	0.59
Roads area ratio	1.446	0.692	1.414	0.707
Forest area ratio	1.568	0.638	1.567	0.638
Daily Maximum Rainfall	1.713	0.584	1.695	0.59

Table 3 shows the result of the multi-collinearity analysis. If the tolerances were less than 0.2 or the VIF was greater than 5, it was interpreted that there was multi-collinearity (O'Brien, 2007). The VIF and tolerances of the factor 'slope' were 5.013 and 0.199, respectively. Since the VIF was greater than 5, and the tolerance was less than 0.2, the 'slope' was removed for LSA. As a result of

performing the multi-collinearity analysis again without the 'slope', it was found that there was no multi-collinearity among the factors.

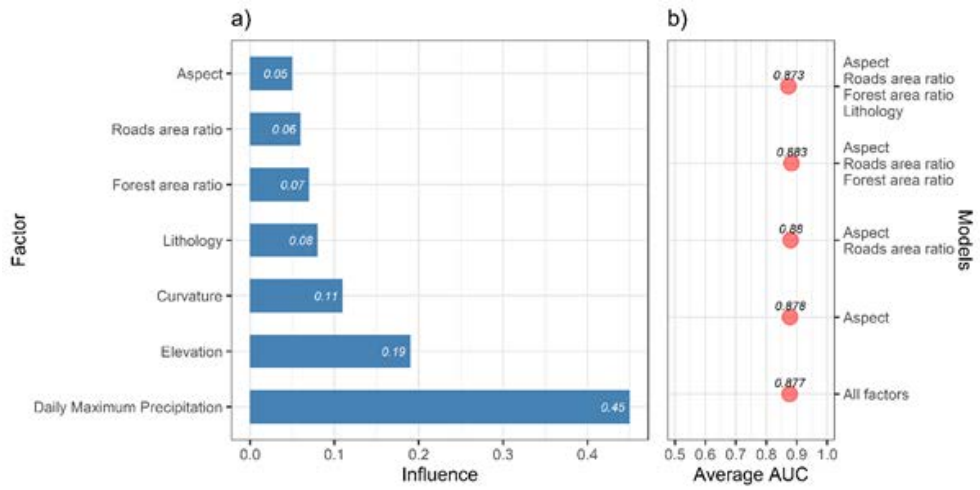

Figure 4. a) Influencing factor analysis, and b) comparison of the average AUC value of the result of LSM by removing the influencing factor analysis result with low influence in order

Figure 4 shows the result of influencing factor analysis and the average AUC value of the result of LSA by removing the influencing factor analysis result for low influence in order. As can be seen from the results (Figure 4a), the most influential variable was 0.45, which was shown as daily maximum precipitation (DMP). The variable with the least influence was aspect (0.05), which was much higher than the altitude (0.023) that most affected the colluvial landslide calculated by Zhou et al (2018). The factor's influence is difficult to compare accurately because the number of variables used is different, but even after taking this into account, it was a high value. As a result of performing LSA ignoring factors having low influence, it can be seen that the AUC value is not significantly different (Figure 4b). Then, LSA was implemented by using all factors without factor 'slope,' which was removed at the multi-collinearity analysis. In addition, DMP appeared to be the most influential factor, similar to

those studies that analyzed flooding in coastal areas (Park & Lee, 2020). These results reiterate the importance of responding to heavy rain as part of disaster management.

3.2 Comparison of machine learning algorithms

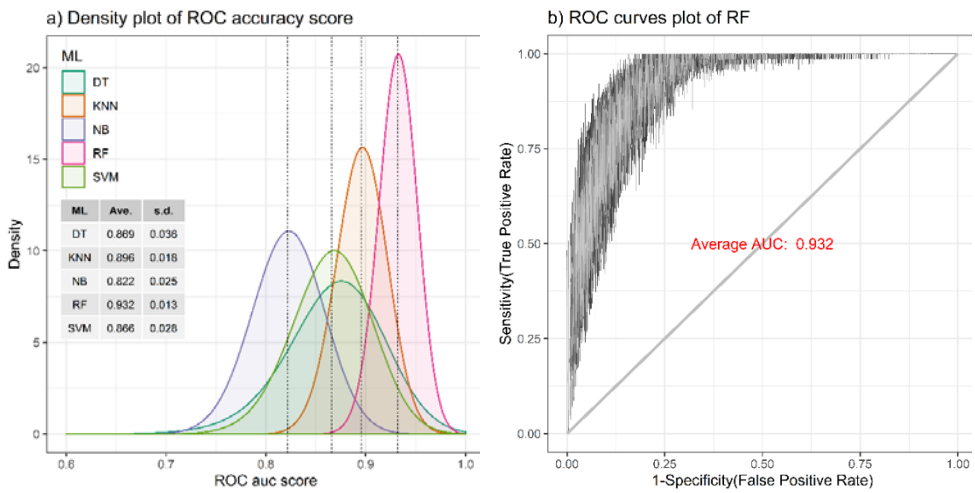

Figure 5. (a) Comparison of the accuracy scores of each ML algorithm and (b) plot of ROC curves of the RF algorithm.

As shown in Figure 5a, the accuracies of the average ROC curves produced from the five algorithms were: 0.822 (NB), 0.896 (kNN), 0.869 (DT), 0.932 (RF), and 0.866 (SVM). The accuracy of the model using RF was the highest of the 5 models; however, the results of the other models were also high. In addition, the graph depicting the density of RF versus ROC accuracies shows that the kurtosis of the graph is slightly higher than those of other algorithms. This indicates that the results obtained with RF were precise. Therefore, we used RF for the landslide susceptibility mapping and prediction. Figure 5b shows the ROC curves obtained using RF.

3.3 Predicting landslide susceptibility

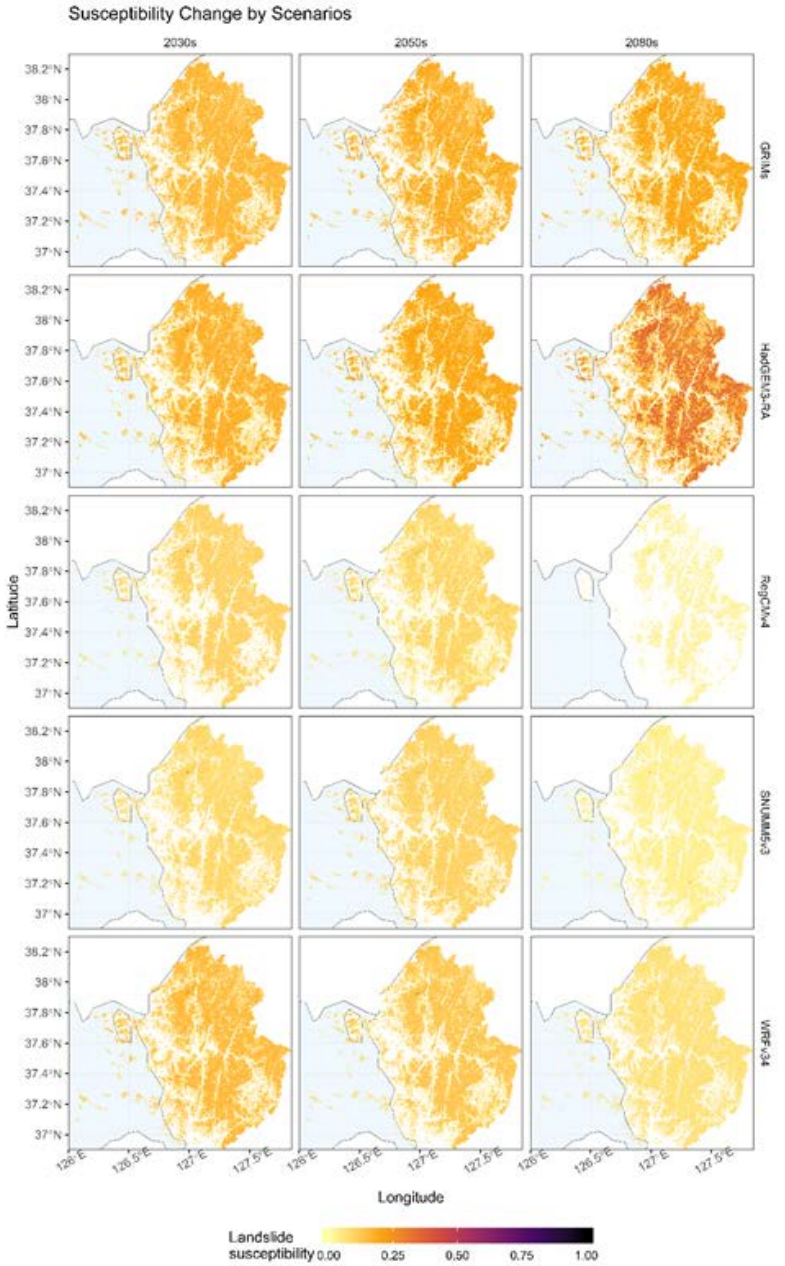

Figure 6. Landslide susceptibility changes for different RCMs under RCP climate change scenario 8.5 from the 2030s to the 2080s.

The predictions were conducted using monthly average rainfall amounts. In this process, rainfall predicted to occur in the

future was used as a density function to consider the uncertainty of future rainfall. The rainfall was estimated by substituting the kernel density in one month, and the model was used to calculate the monthly predicted susceptibility. By comparing the probability of monthly susceptibilities, we observed that the probability increased during the months of June, July, and August for most of the RCP 8.5 scenarios. Based on these data, the average probabilities of susceptibility for each scenario (8.5, 2030s, 2050s, and 2080s) during June, July, and August were calculated to create a probability map of the maximum susceptibility for the given scenario.

The results of the predictions are shown in Figure 6. These results were obtained for the RCP climate change scenario 8.5 using five different RCMs with the 2030s, 2050s, and 2080s as the target periods. The susceptibility was high, on average, when HadGEM3-RA was used. Moreover, when GRIMs were used, the susceptibility increased over time, while the remaining three RCMs had tendencies for reduced susceptibility. This is because the peak precipitation value appeared differently depending on the scenario considered by the RCMs (Figure 3i).

4. Discussion

4.1 Analysis of results from different ML algorithms

This study accounted for the model uncertainty by performing LSA using five different ML algorithms. The resulting performance was good for most models, with RF having the best performance, similar to the findings of previous studies such as those using supervised learning classification with spatially geographical

data (Cracknell & Reading, 2014; Chen et al., 2017b; Naghibi et al., 2017; Pourghasemi et al., 2020).

4.2 Difference in susceptibilities based on land cover type

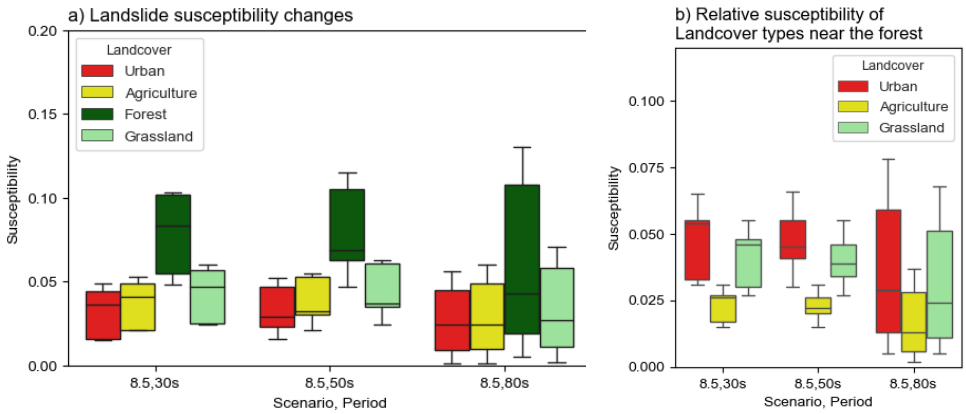

Figure 7. a) Changes in landslide susceptibility according to land-cover type in the RCP 8.5 scenario for different periods. b) Relative susceptibility of other land-cover types near forest areas.

We determined how susceptibility changed according to the type of land cover when the same land cover is maintained in the future using five different RCMs under climate change scenario 8.5 (Figure 7a). The risk in forest areas is largely due to the occurrence of landslides. The uncertainty of the results increased over time. This can be attributed to the uncertainty of climate models and the distribution characteristics of the precipitation values in each RCM.

Moreover, it is important to know which areas among other land cover types near the forest are more sensitive in the future. Figure 7b shows how three land-cover types near the forest are relatively susceptible during the target periods under the RCP 8.5 scenario. The uncertainty increased with time; however, the urban area remained more susceptible than other areas. This is because

many urban areas are distributed around forest areas that have high susceptibilities to landslides. More indirectly, it could be explained that due to the high urbanization around the forest area, economic damage is high when landslides occur due to heavy rains. These results highlight the need for future efforts in land-use planning to reduce landslide susceptibility in urban areas located near forest areas.

5. Conclusions

This study evaluated landslide susceptibility in the metropolitan area that includes Seoul, South Korea. Prior to LSA, multi-collinearity was analyzed among eight factors. As a result of the multi-collinearity analysis, the factor 'slope' had high multi-collinearity, and it was removed for LSA. Then, to improve the ML model's performance, the influencing factors were addressed by using the IGR technique. The most influential factor was the daily maximum precipitation (0.45). In addition, the results of performing LSA while ignoring factors having low influence in turn showed that the AUC value is not significantly different from each result. Based on evaluations using five different ML algorithms with seven factors, the average AUC using RF exhibited the best performance (0.932). To predict future landslide susceptibility, projected future precipitation values from the RCP climate change scenario were used, and the target years were set to the 2030s, 2050s, and 2080s. Additionally, various future possibilities were predicted based on five different RCMs. This allowed us to consider future uncertainties. The landslide susceptibility generally increased over time, with the

exception of some results where RegCM4, SNURCM, and WRF were used.

From this study, we arrived at the following additional inferences. First, among the five ML algorithms used to assess landslide susceptibility, RF had the best performance. Similar studies using supervised learning classification have shown comparable results. Second, we found that it is necessary to reduce the landslide susceptibility in forest areas, based on an analysis of changes in landslide susceptibility over the target period (2030s, 2050s, and 2080s) and land cover types (urban, agricultural, forest, and grassland). We found that urban areas were more susceptible than other land-cover types because they were distributed around the forest areas that were estimated as more susceptible to landslides. It might indicate that many areas were urbanized around the forest area, and it also highlighted the need for future efforts in land-use planning to reduce landslide susceptibility.

Additionally, the result and process of this study reveal some limitations. First, since the purpose of this study was to predict future landslide susceptibility, the effect of data resolution was not considered, although the resolution of the area of the target site was larger than the resolution of the two other studies (Dou et al., 2015; Chang et al., 2019); thus, the resolution of data should be considered in future studies because it is important for determining landslide susceptibility. Second, we predicted future landslides by assuming that the socioeconomic factors (as indicated by land cover type) did not change. In the future, socioeconomic factors may become more important for determining susceptibility; therefore, these factors should be considered in future studies.

Chapter 4. Adaptation strategies to future coastal flooding: performance evaluation of green and grey infrastructure in South Korea

1. Introduction

Climate change is one of the most dangerous environmental problems that humans face at present and poses profound global and regional impacts. Climate change contributes to an increase in the frequency and intensity of extreme drought, heavy rain, and heat waves (IPCC, 2022). Floods caused by heavy rains and abnormal climates cause substantial damage to various regions around the world. (Merz et al. 2021). It is expected that more than a billion people will live in low-lying coastal areas by 2060 (Neumann et al., 2015). Coastal areas with a high concentration of economic activity are particularly at risk of extreme weather events owing to climate change, as they are exposed to cyclones, tsunamis, and other coastal hazards. (Kron, 2013; Reguero et al., 2015; Ferreira et al., 2019; Reguero et al., 2020). The global sea level is currently rising by 3–4 mm per year because of ocean warming and land ice melting and is projected to rise by 0.3–2.0 m by 2100 (Watson et al., 2015; Yi, et al., 2015). Therefore, the damage to coastal areas will likely be even more severe in the future because of the combined risk effects of climate change, including heavy rain and sea-level rise (Barnard et al. 2015; Caldwell et al. 2009; Church et al., 2013; Vitousek et al., 2017).

Korea is a peninsula surrounded by the sea on three sides, with many large cities being in coastal areas. Approximately 27.5% of the total population of Korea lives in coastal areas (Oh et al., 2020)

and 1.3 million people are forecasted to experience coastal flooding by 2050, and further exposed to Pacific typhoons (Kulp & Strauss, 2019). The cost of damage to property caused by Typhoon Maemi (2003) and Rusa (2002) was $6.6 billion and $5.3 billion, respectively (Ministry of Public Safety and Security, 2015). Without an appropriate response to coastal flooding, increasing property loss and fatalities are inevitable.

The steps for responding to disaster events, such as coastal flooding, can be divided into three major parts (Reguero et al., 2018): 1) identifying areas vulnerable to hazards, 2) predicting how much damage will be caused by the risk, and 3) taking appropriate measures to reduce the risk and damage. All three processes are integral, but the last step must consider the efficacy of different strategies. The best possible coastal management measures considering cost and effectiveness will be essential in the future to optimally use limited resources (Ferreira et al., 2019). These efforts can proactively facilitate risk mitigation before a disaster strike (FEMA, 2018; Reguero et al., 2020).

Past studies have investigated the effectiveness of measures and strategies to reduce the risk of coastal flooding. Reguero et al. (2018) quantitatively evaluated the cost-effectiveness of natural-based solutions (NBS) and structural techniques by comparing the effects of reducing the risk of coastal flooding in the Gulf Coast region at the county level. They found that applying NBS has a risk-reduction effect of 3.5 times compared to the investment. Ferreira et al. (2019) used a numerical model to compare the performance of beach nourishment with shelter removal and movement for reducing the impact of storms in 55-km coastal areas. It was found that using

both options together or nourishment alone could avoid almost all observed effects. Vousdoukas et al. (2020) considered composite events with respect to climate and developed analytical tools to optimize adaptation strategies (ASs) and assess their impacts during the present century. They revealed that the use of dykes could reduce the economic effect of European flood damage by 23.7–32.1%. Tiggeloven et al. (2020) proposed a framework that considered different flood risk factors to assess the future benefits and costs of structural protection measures related to coastal flooding on a global scale at the country level. The benefits of the application of dykes were calculated according to four objectives, including maintaining the current protection level, and all showed positive effects. Creach et al. (2020) compared the implementation cost and efficiency of different housing ASs when exposed to flood risk to mitigate the level of risk in La Guérinière, France. Relocation reduced the risk by 100%, but was expensive while protection (dykes and seawall) reduced risk by 26–46% with an investment of 3.8 billion €. Therefore, numerous studies have been conducted to compare and evaluate the effectiveness of strategies for mitigating and reducing the risk of coastal flooding, although with different spatiotemporal extents. However, previous studies often did not consider spatiotemporal downscaled analysis. A study of the smallest spatial units was carried out by Reguero et al. (2018) at the county level. Additionally, a predictive study using future climate change scenarios (Tiggeloven et al., 2020) at the country level did not consider downscaled spatial range. The impact of and vulnerability to climate change may vary owing to the spatiotemporal characteristics of nations, under both current and future climate conditions (IPCC, 2007, 2014, 2022). It is

thus necessary to consider spatiotemporal downscaled analysis to compare and evaluate the effectiveness of ASs in mitigating the risk of coastal flooding.

Therefore, this study aimed to explore the risk of coastal flooding driven by heavy rain and sea-level rise through detailed spatial and temporal modeling. To answer the question of how ASs for coastal flooding will be effective in the face of climate change, we examined the effects of green space and seawalls according to the geographic locations and urbanization intensities of cities along the South Korean shorelines.

2. Materials and Method

2.1 Study area

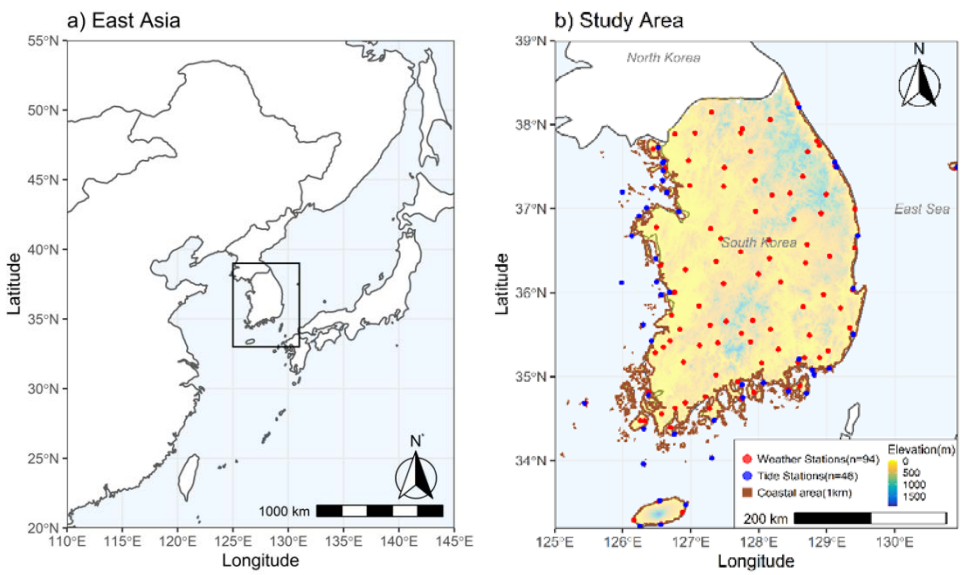

Figure 1. (a) Countries in East Asia, (b) South Korea and the study area (Coastal area: 1 km from the coastline, resolution:0.25km, Geodetic Datum: WGS84), Adapted from "Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms" by Park & Lee, 2020, Environmental Research

The study sites included the coastal areas of South Korea (33–38 °N, 125–131 °E). Following the “coastal management law” in the country, the spatial scope of this study was set to a 1 km inland buffer area from the coastline at a resolution of 250 m. The total length of the coastline of South Korea is 14,962.8 km and Figure 1 shows the locations of 68 weather stations and 46 tide stations in the nation.

For the past 106 years (1912–2017), the average annual temperature in South Korea has been 13.2 ℃ and the annual precipitation has been 1,237.4 mm (National Institute of Meteorological Science, 2018). In general, summer precipitation (June–August) is 710.9 mm, accounting for 54% of the annual precipitation. In addition, over the past 30 years, the amount of precipitation during summer has increased by an average of 0.11 mm per year. The frequency and intensity of extreme rainfall on the Korean Peninsula have also been gradually increasing since the mid-1990s (Korea Meteorological Administration, 2020).

2.2 Data

To analyze the risk of coastal flooding, a total of six variables, including the flood trace, were used with reference to Park and Lee (2020) (see Table 1). These variables have diverse data types and resolutions and were refined to match a resolution of 250 m. In addition, because traces of flooding existed from 2003 to 2018 and the associated climatic and environmental conditions on the days

when flooding occurred were different for each event, a refinement process was performed to match the event-based time-series data. For example, in the case of flooding that occurred on September 1, 2003, the daily maximum precipitation and mean tide data for that date were coordinated.

Table 1. Data list with source and period

Category	Variable (unit)	Period	Data Type	Source
Marine	Mean Tide (mm)	2003-2018	Point	*KHOA
Meteorological	Daily Maximum Precipitation (mm)			**KMA
Geophysical	Elevation (m) Slope (degree)	2010	Grid (250m)	***ME
	Urban area (%)	2000, 2009	Polygon	
Coastal Flood Trace		2003-2018	Polygon	****LX
Adaptation Strategies (ASs)	Seawall (m)	2020	Polyline	
	Greenspace (%)	2013	Polygon	***ME

*KHOA: Korea Hydrographic and Oceanographic Agency (http://www.khoa.go.kr)
**KMA: Korea Meteorological Administration (https://www.weather.go.kr)
***ME: Ministry of Environment (http://me.go.kr)
****LX: Korea Land and Geospatial Informatrix Corporation (https://www.lx.or.kr)

2.3 Comparison of machine learning (ML) techniques and coastal flooding risk analysis

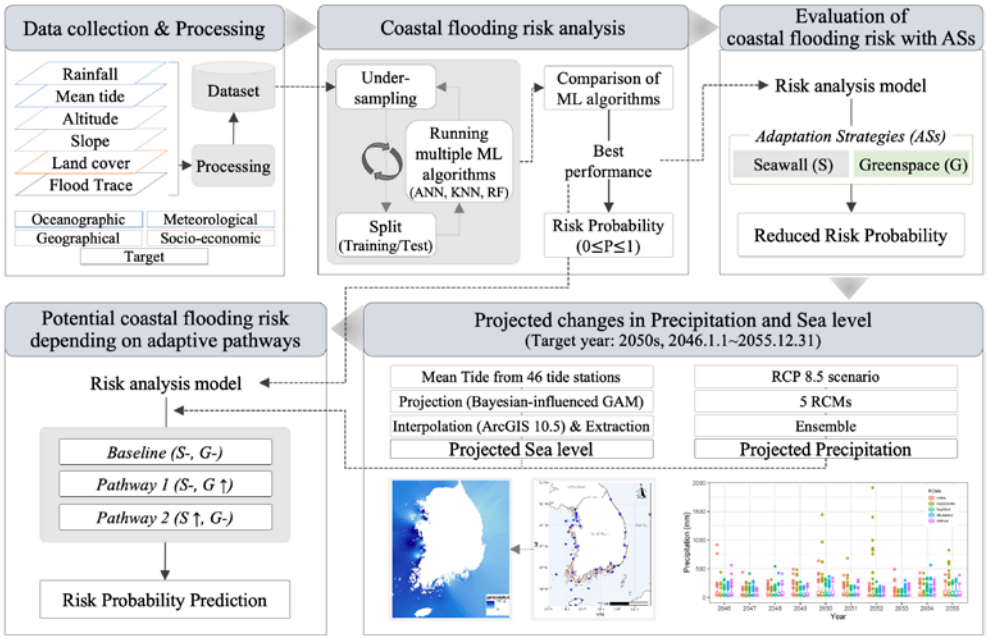

Figure 2. Research flow

As shown in Figure 2, to predict coastal flooding potentials, the under-sampled dataset was divided into training (70%) and test (30 %) sets and analyzed using three different machine learning (ML) techniques: k-nearest neighbor (kNN), random forest (RF), and artificial neural network (ANN). The reason for under-sampling is that coastal flooding trace data were imbalanced, with more non-occurrence than occurrence. Through more than 5,000 iterations of training per technique, the average performance obtained for the techniques was compared. Coastal flood risk was analyzed using the technique showing the best performance. The technique with the best performance was selected by comparing the Area under the Receiver operating characteristic (ROC) curve (AUC), which is mainly used to evaluate the performance of machine learning (Huang & Ling, 2005).

The probability results ranged from 0 to 1, with a higher probability value indicating a higher risk of flooding.

kNN, proposed by Cover and Hart (1967), is a supervised learning ML algorithm that is simple and easy to implement. The setting of the k value that indicates proximity has a significant influence on the results of the algorithm (Bhavsar & Ganatra, 2012; Kim et al., 2012). The RF algorithm, described by Breiman (2001), is an ensemble learning method that constructs and operates several decision trees during the training period and is widely used in research along with ML techniques that use neural networks. ANNs are computational models inspired by the human brain (Yang, 2008). ANNs are used to model complex relationships between input and output information through a network of interconnected nodes to find patterns in the data. (Potdar & Kinnerkar, 2016).

2.4 Evaluation of coastal flooding risk with ASs

Using the ML technique that showed the highest performance, the effect of reducing the risk probability was computed under AS scenarios. As shown in Table 1, the seawall, representing the gray infrastructure installed to protect coastal areas, and the green space, representing the green infrastructure, were compared. Green space and seawalls play an important role in the protection of coastal flooding (Dong et al., 2020; Jeong et al., 2021). In this study, green spaces were artificially created and developed grasslands, including urban parks, open lawns, and riparian buffers. In previous studies, green infrastructure was found to be effective at reducing runoff by 3–47%. (Zhang et al., 2012; Ahiablame & Shakya, 2016; Arjenaki et al., 2021; Li et al., 2021). Therefore, the same reduction range was

applied in this study to predict the potential of green spaces to mitigate surface flooding. Meanwhile, in the case of the seawalls, which are mainly installed along the coastline and serve to protect the land, a range of elevations was applied, elevating the height of the seawall, and the associated slope was calculated and updated in the prediction model.

2.5 Potential coastal flooding risk depending on different adaptive pathways

Table 2. Pathways according to the degree of application of adaptation strategy (Target of the period: RCP 8.5, 2050s)

Adaptive Pathways	Application of adaptation strategy	Degrees of application					
Baseline	No adaptation	–					
Pathway 1	Greenspace	5%	10%	15%	20%	25%	30%
Pathway 2	Seawall	1.5m	3m	4.5m	6m	7.5m	9m

Among the variables used in the analysis, daily maximum precipitation and mean tide are time-variant factors; therefore, future values were projected based on carbon emission scenarios developed by the Intergovernmental Panel on Climate Change (IPCC). Using their fifth assessment report's RCP 8.5, which represents the "business as usual" scenario with a high-energy demand, a long-term precipitation trajectory was predicted by the middle of the century. We used five regional climate models (RCMs) provided by the Regional Climate Detailing Project in East Asia (CORDEX-EA: Coordinated Regional Downscaling Experiment, East Asia; source: http://cordex-ea.climate.go.kr/cordex): CCLM, HadGEM3-RA,

RegCM4, SNU-RCM, and WRF. Ensemble methods using multiple climate models allow researchers to consider the uncertainty of different models (Parker, 2013). Time-series tide data from 46 stations located along the coast of Korea were used. As the observed mean value approximates a sine distribution, a Bayesian-influenced generalized additive model (GAM) was used to predict the future value for each station. The longitudinal data were then spatially interpolated onto the grid block for the input of the prediction model. The model was used to calculate the potential coastal flooding risk by extracting the space corresponding to the study site from interpolated data (see Figure 2). In this study, we analyzed the potential coastal flooding risk probability, setting the year 2050 as the target. We selected 2050 because this is the year of the global target to achieve "net zero", meaning that net carbon emissions should be zero (IPCC, 2018).

Under the forecasted future climate conditions, two adaptive pathways were developed based on the flood mitigation techniques adopted and multiple scenarios were formulated for each pathway according to the intensity of application (see Table 2). 1) The baseline pathway maintained the current level of ASs with no additional implementation. Pathway 1 hypothesized that green spaces were applied to 5–30% of the entire study area at 5% increments on a cell basis. Pathway 2 applied seawalls with heights ranging from 1.5 to 9 m at 1.5 m increments. These application levels were determined using the standard construction cost per unit area for each technique. Regarding green space, the standard cost was 73,000 ₩ ($66.4 US dollars) per 1 m² according to the "national land planning and utilization act" in South Korea. For seawall, the cost was 500€

($550 US dollars) per 1 m in height and length (Creach et al. 2020). In other words, 1% of the area of green space equated to 625 m² (considering the resolution of the analysis), with corresponding construction costs of $41,500. With the same amount of money, a seawall with an approximate length of 75m of 1 m height (=$41,500/$550) can be built. Thus, the construction costs of 75 × 1 m (length × height) of seawall were equivalent to that of a 250 × 0.3 m seawall in a unit cell. That is, a 1% increase in green space corresponded to a 0.3 m increase in seawall height (i.e., 5% is equal to 1.5 m of seawall).

3. Results

3.1 Performances of ML algorithms

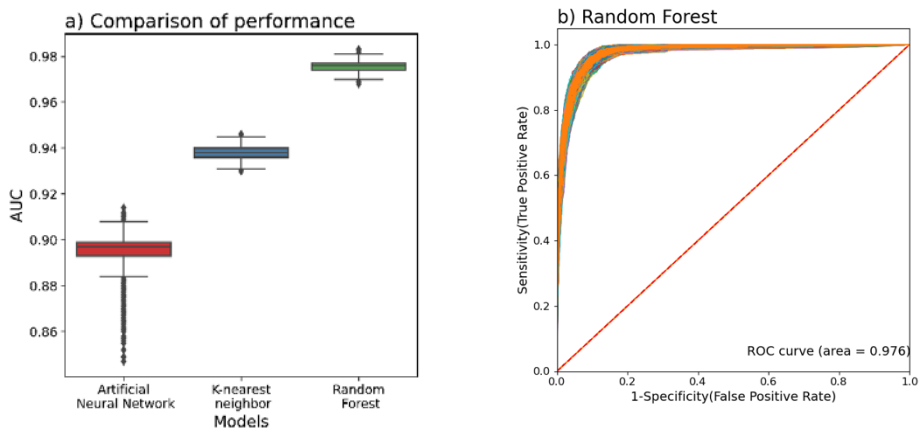

Figure 3. Comparison of machine learning algorithms' performances (a: comparison of performances among 3 different machine learning techniques, b: ROC curve plot by using Random Forest)

According to the results obtained through approximately 5,000 runs using three different ML algorithms, the RF model performed the best, with an average AUC of 0.976, followed by kNN (0.938) and ANN (0.896) (Figure 3a). Although the differences were

not significant, the dispersion of AUC was the lowest when RF was used, implying that RF was the most stable model. Figure 3b shows the ROC curve plot obtained using RF. The ROC curve is suitable to indicate the reliability of the model, and a more curvilinear shape towards the left or top indicates a higher reliability of the model.

3.2 Coastal flooding risk with ASs

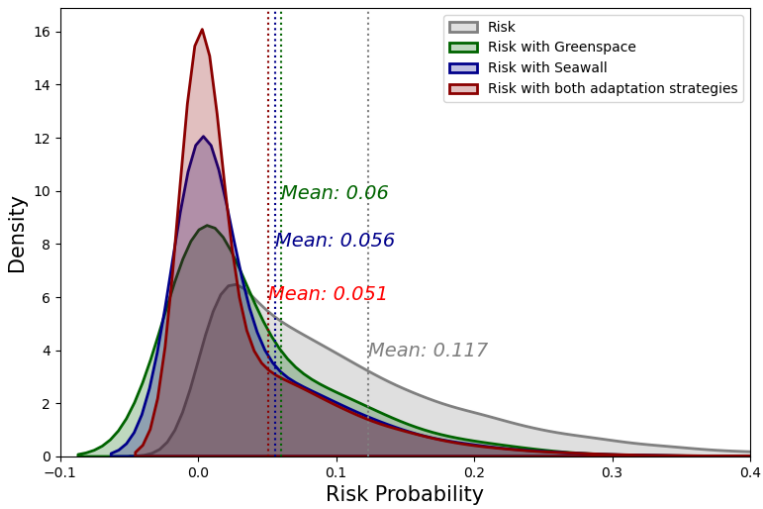

Figure 4. Comparison of coastal flooding risk probability depending on consideration of two adaptation strategies under current level

The current level of coastal flood risk was probability calculated using the best-performing algorithm, RF, and the change in risk probability was analyzed when two spatially distributed ASs were applied (Figure 4). The average risk probability was 0.117 in the absence of ASs. When green spaces and seawalls were applied at the current level, the average risk probabilities were 0.066 and 0.056, respectively. Finally, when both were applied together, the average risk probability was 0.051. There was no significant

difference in the average risk probability between the two strategies, applied separately or together.

Despite the insignificant reduction rates of the strategies applied, the spatial analysis demonstrated their varying performances by geographic location. Figures 5 a, b, and c show the mapping results of the coastal flood risk probability when AS was not applied (a), when AS was applied (b), and when the risk probability was reduced by more than 0.9 (c). The blue dots in Figure 5 c represent the points where coastal flooding occurred, as shown in Table 1. Although the average reduction in risk probability may seem low for the entire considered area (below 1%), regions disproportionally benefitted from AS; areas where the actual flood occurred particularly experienced a greater reduction in risk probability than other areas. Thus, it is important to use an adaptation strategy to reduce coastal flooding risk.

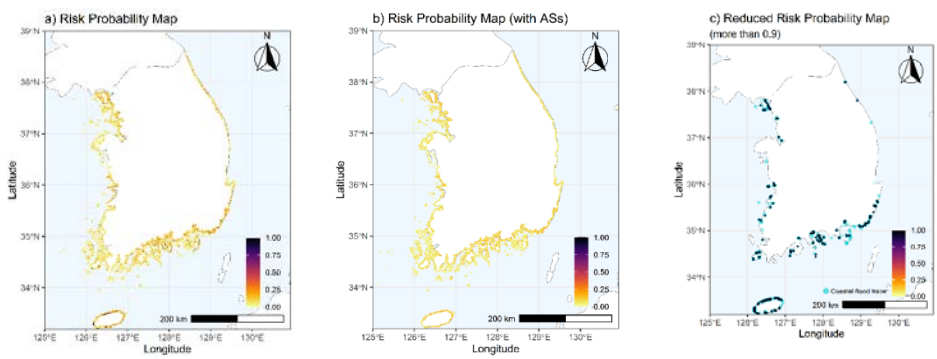

Figure 5. Comparison of coastal flooding risk probability spatially depending on consideration of adaptation strategies under current level (a: without adaptation strategies, b: with both adaptation strategies, c: reduced to a probability of more than 0.9, blue dots are the points where coastal flooding occurred from 2003 to 2018)

3.3 Potential coastal flooding risk according to different adaptive pathways

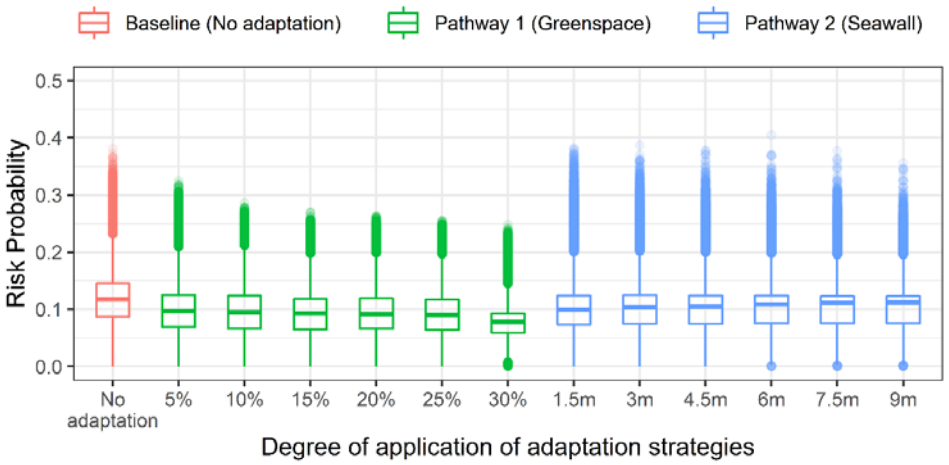

Figure 6. Coastal flooding risk probability according to the degree of application of each adaptation strategy in the three adaptive pathways (RCP 8.5, 2050s)

As shown in Figure 6, the results of applying varying intensities of ASs in the two pathways revealed that green spaces generally have a better reduction effect than seawalls. The effect of green spaces improved slightly as the coverage area increased, and the effect was noticeable when it was expanded from 25% to 30%. However, for seawalls, the average risk probability did not change significantly when the height was increased, even though the frequency of high extreme values tended to gradually decrease. Ultimately, seawalls were less effective than green space, but they reduced the degree of risk to some extent for high-risk areas.

Nevertheless, seawalls are an effective and economical structural strategy that are primarily considered for reducing the risk of coastal disasters (Duvat, 2013) and can be used in combination with green infrastructure when federal and local budgets are restrained. A prioritized approach should be taken, implementing

green spaces in the most vulnerable areas while applying seawalls as supplemental systems in moderate- or low-risk areas.

Next, to identify the spatial differences in risk probability, the spatial distribution according to the AS application scenario, targeting 2050, was compared for five different RCMs within the RCP 8.5 climate change scenario to determine the probability of potential coastal flood risk (Figure 7). When no ASs were applied until 2050, HadGEM3RA and GRIMs showed a higher coastal flooding risk than the other three RCMs. This is because the result of the projected precipitation slightly differed according to the difference in the statistical techniques with which the RCMs are created (Déqué, 2007). Therefore, to consider the uncertainty of each model, the results of various models were assembled and the combined results were compared. As shown in Figure 6. considering the ensemble of baseline, pathways 1, and 2, green space (30%) application showed a greatly reduced risk spatial distribution than seawall (9 m) application.

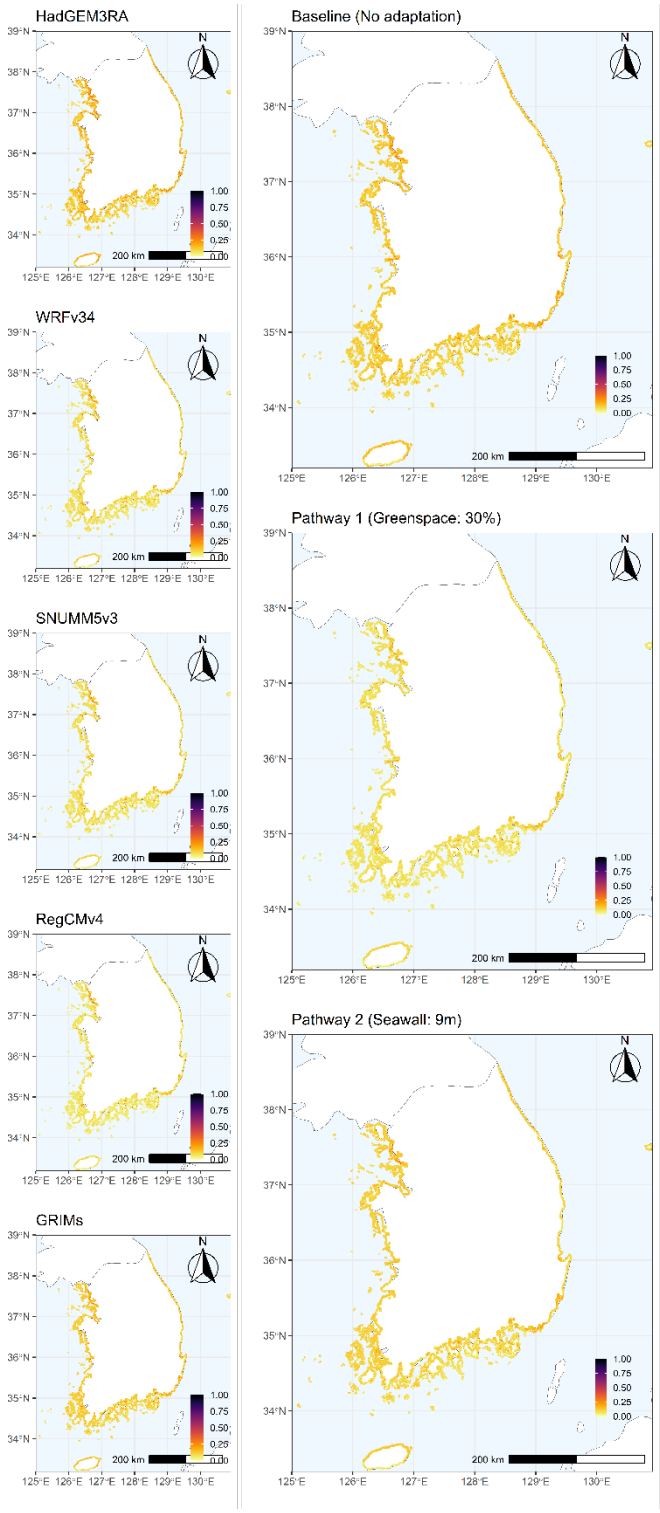

Figure 7. Spatial distribution of coastal flooding risk probability using five different RCMs at 2050s of RCP 8.5 (Left: comparison of five RCMs without adaptation strategy, Right: comparison of ensemble with each adaptation strategy)

4. Discussion

4.1 Effect of AS according to spatial characteristics

Considering that the general opinion in climate change research is that the impact and vulnerability of climate change can vary greatly depending on the regional characteristics of climate change adaptation (IPCC, 2007, 2013, 2020), it is important to identify the differences in risk and vulnerability according to spatial characteristics. Therefore, a comparison of risk probabilities according to geographic and environmental differences was conducted to determine the differences in the application of ASs.

4.1.1 Differences in risk according to coastal geographical characteristics

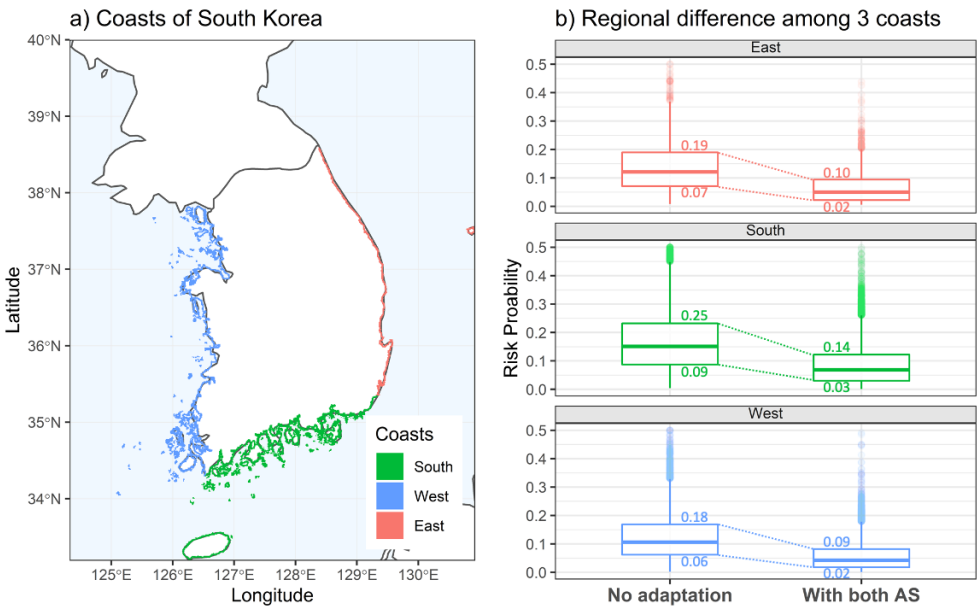

Figure 8. Regional difference among three coasts (a: three coasts of South Korea, b: coastal flooding risk probability among 3 coasts without and with two adaptation strategies), Adapted from "Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms" by Park & Lee,

In the case of South Korea, a peninsula-shaped country, the geographical characteristics of the east, west, and south coasts differ. The east coast has a monotonous coastline with a deep-water depth, while the west coast has substantial difference in tides owing to the low land and water depth. On the southern coast, there are numerous large and small islands with complex coastlines. (Park & Lee, 2020) which respond differently to flooding (see Figure 8). When AS was not applied, the risk probability of the southern coastal region was found to be the highest, on average. The probability between the 25th and 75th percentiles ranged from 0.09 to 0.25, a wider distribution than that for the other coasts. The southern coastline is characterized by an irregular pattern of low-lying lands and a high density of islands (see Figure 8a). The region is the most vulnerable to flooding when tropical storms pass over the south coast during the typhoon season and reach landfall from June to September. However, when two ASs (green space and seawall) were applied, the probability decreased, with comparable interquartile ranges across the three regions. It is important to note that the high extremes of risk probability significantly reduced on applying ASs, demonstrating their importance.

4.1.2 Differences in risk by urbanization rate

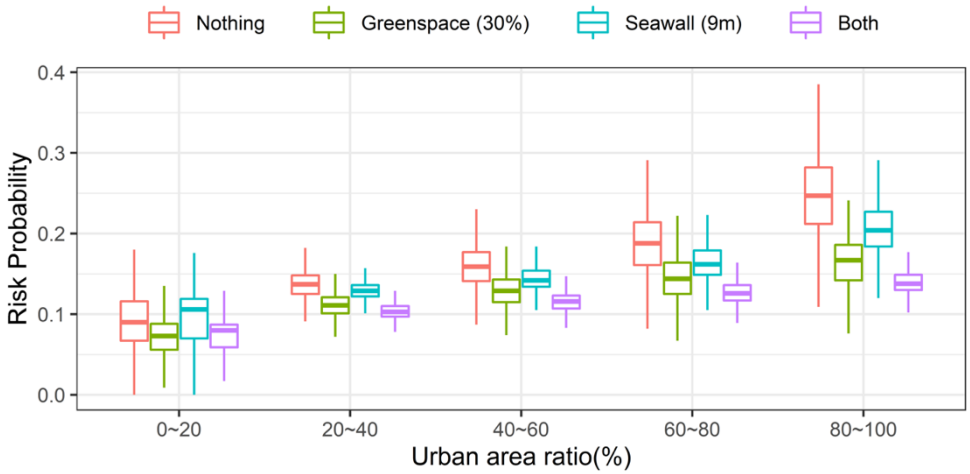

Figure 9. Coastal flooding risk probability according to the urbanized area ratio change and whether adaptation strategy is applied

Figure 9 shows the change in the distribution of risk probability according to the urbanization ratio and AS type. Except for regions where the urban area ratio was less than 20%, the risk probability increased with the urban area ratio regardless of whether an AS was applied. It should be noted that when AS was not applied, the risk probability increased substantially as the urban area increased. On average, the risk probability with more than 80% urbanization was 2.5 times greater than that with 0 to 20% urbanization. However, this increase can be mitigated by the application of ASs. Comparatively, when comparing the risk probabilities depending on whether two ASs were applied, if only one was applied, green spaces had a greater mitigation effect than seawalls. Importantly, when these two techniques were applied together, the reduction effect increased as the ratio of the urban area increased; the slope of the baseline scenario (red) was steeper than that of the combined strategy (purple) and the difference between

them increased as the area of an urban area increased. This indicates the importance of investments in ASs in highly urbanized areas.

In summary, Figure 9 implies that the risk in urban areas with high population density and socioeconomic activity may increase in the future. According to the forecast that more than one billion people will live in low-lying coastal zones by 2050 (IPCC, 2019), it is imperative to prepare for this.

4.2 Importance of nature-based solutions as ASs

To respond to the climate crisis, efforts by the international community to achieve carbon neutrality to not exceed a global temperature increase of 2 ℃ are being accelerated. However, despite the aim to achieve a carbon-neutral society, the increasing frequency and intensity of natural disasters due to rising global temperatures driven by emitted and accumulated carbon will continue and efforts must be made to adapt and respond to these. Therefore, climate change ASs are as important as reducing and absorbing carbon emissions.

NBSs have been emphasized for achieving goals in international communities, such as the Sendai Framework for Disaster Risk Reduction, Sustainable Development Goals (SDGs), and Paris Climate Agreement (Reguero et al., 2020). In particular, the importance and necessity of ecosystem-based ASs, such as green or natural infrastructure, have been emphasized among potential strategies to reduce flooding risk in coastal areas (Cheong et al., 2013; Spalding et al., 2014; World Bank, 2018; Jongman, 2018; Reguero et al. 2018; 2020; Ferreira et al. 2019; Jeong et al., 2021),

which is consistent with the findings of this study. In addition, NBSs can play an important role in realizing carbon neutrality. As they are based on nature, NBSs include ecological elements, such as vegetation and soil, which have the function of absorbing carbon (IUCN, 2012). Therefore, investments in NBSs are of utmost importance for increasing the resilience of urban spaces to climate change.

5. Conclusion

This study predicts areas at high risk of flooding in coastal regions of South Korea because of climate change, which will likely cause extreme rainfall and sea-level rise, using spatial and temporal downscaling of future scenario values. Changes in risk probability according to the application or non-application of ASs were confirmed. To calculate the coastal flooding risk probability, the RF, which showed excellent performance, was utilized by comparing the performances of three ML techniques (RF, ANN, and kNN). To predict the potential coastal flooding risk probability in the future, the precipitation data were projected through an ensemble of five RCMs of the RCP 8.5 scenario provided by IPCC AR5. In addition, using real-time tide data from 46 tide stations on the coast of South Korea, the spatial and temporal data were projected by the Bayesian-influenced GAM.

In this study, “nature” based green spaces and “gray-infrastructure” based seawalls, which are representative structural approaches to coastal flooding, were used as ASs. There was a clear difference in the risk probability distribution when ASs were applied

or not applied and the effects of green space and seawall did not differ significantly (see Figure 4). However, the effects of the two ASs on the future potential risk probability were somewhat different, and when 2050 was predicted as a target, green spaces were slightly more effective than seawalls (see Figure 6).

However, this study had a limitation in that it did not utilize more diverse ASs. Several structural measures have been proposed to reduce the risk of coastal flooding, including sand nourishment, reefs, and dykes (Singhvi et al., 2022). Nevertheless, the importance of ecosystem based NBS strategies was confirmed by utilizing the two contrasting strategies. In addition, this study only considered the effects of rainfall and sea-level changes as causes of coastal flooding. As in previous studies that considered coastal flooding through the interaction between rainfall and sea level (Eilander et al., 2020; Park & Lee, 2020), in coastal areas, like inland areas, inundation caused by heavy rains is just as important as flooding by storms and waves.

However, this study has several notable aspects. The impact of climate change differed spatially and different adaptation measures are required according to the characteristics of the region. Therefore, it is necessary to consider the characteristic differences of the coasts of Korea because there are differences in the geographical characteristics of the east, west, and south coasts and between urban and non-urban areas. As a result of comparing the coastal flooding risk probability of the east and west coasts of Korea, the risk probability of the south coast was slightly greater than that of the east and west coasts, with a comparatively wider distribution. This may be because of the geographical characteristics of the southern coast. Moreover, the risk probability increased as the urbanization

rate increased. This means that a climate change response strategy is necessary for coastal cities, considering that the population and socioeconomic activities will increase in urban areas in the future.

Chapter 5. Conclusion

Climate change is an urgent threat to our generation. Natural hazards have become more unpredictable, occurring more frequently and with greater force, due to climate change (Berz et al., 2001; Kundzewicz et al., 2014; UNISDR, 2015). Natural disasters in Korea are mostly caused by meteorological events, damage due to hazards such as flooding, landslide in Korea will increase further due to climate change in the future. Therefore, risk management, which analyzes and evaluates hazard risk related with heavy rainfall such as flooding and landslides, is needed to prepare for the long term. Also, it is very important what measures should be taken to reduce risk and damage in response to natural disasters caused by climate change.

Firstly, South Korea is a peninsula with several large cities situated along the coast and 27.5% of its total population living in coastal areas (Oh et al., 2020). Thus, it will be more vulnerable in the future to predicted climate change impacts such as sea level rise and extreme weather events. We evaluated the future probability of coastal flooding events–based ML algorithm, namely kNN by using six variables. A risk probability map was developed, and to evaluate future coastal flood risks due to climate change, we estimated the future risk probability using forecasted tidal and future rainfall. As a result, the risk probability increased over time and the risk probability increased in the southern coastal areas more so than in the eastern or western coastal areas. In this study, there are three significant implications: 1) kNN classifier were performed slightly better than the other methods (RF, SVM), 2) rainfall was identified

as a significant factor in this study, 3) future coastal flood risk analysis was analyzed using ensemble methods from different RCMs for considering future uncertainty.

Along with coastal flooding, landslides are a major cause of serious damage to life and property worldwide (Malamud et al., 2004; Gomez & Kavzoglu, 2005; Lee & Pradhan, 2007; Garcia-Rodriguez et al., 2008; Yilmaz, 2010; Pham et al., 2020). In the future, when the impacts of climate change become more severe, sudden heavy rains could cause more damage due to landslides and flooding (Yılmaz, 2009). We evaluated landslide susceptibility in the metropolitan area that includes Seoul, South Korea. Through the multi-collinearity, the factor 'slope' was removed for landslide susceptibility assessment (LSA) because of high multi-collinearity. The LSA was analyzed using five different ML algorithms, then to predict future landslide susceptibility, we used five different RCMs for considering future uncertainties. The landslide susceptibility generally increased over time. In this study, there are three significant implications: 1) To assess landslide susceptibility, RF had the best performance. 2) it is necessary to reduce the landslide susceptibility in forest areas, based on an analysis of changes in landslide susceptibility over the target period (2030s, 2050s, and 2080s) and land cover types (urban, agricultural, forest, and grassland), 3) urban areas were more susceptible than other land-cover types because they were distributed around the forest areas that were estimated as more susceptible to landslides.

The application of an appropriate adaptation strategy to the potential damage of coastal inundation that is further aggravated by the effects of climate change is a very important part of integrated

coastal management. Using machine learning techniques, predict areas with high potential for coastal flooding in Korea's coastal waters due to future climate changes such as extreme rainfall and sea level rise, and check changes in coastal flooding risk probability due to application of adaptation strategies (green areas and breakwaters) did Four important findings from the process were as follows. 1) There was a clear difference in the risk probability distribution when the adaptation strategy was not applied and when it was applied. 2) The effectiveness of reducing potential future flooding risk depends on the choice of adaptation strategy, geographical characteristics, and degree of urbanization. It shows that green spaces are slightly more effective than seawalls when forecasting the future 2050. This indicates the importance of nature-based strategies highlighted in previous studies. 3) It emphasizes the need to prepare adaptation measures according to regional characteristics to the effects of climate change. In the case of Korea, it has independent geophysical and climatic characteristics because it is surrounded by the sea on three sides. Among the three sides, the southern coast was found to have a higher risk of coastal flooding than the east and west coasts. 4) The higher the urbanization rate, the higher the risk probability. This suggests that a climate change response strategy for coastal cities is necessary because the population and socioeconomic activities of coastal cities will increase in the future.

Summarizing the three studies, the risks of disasters that may occur due to the complex effects of climate change were analyzed and strategies for reducing the risks were evaluated. As a method commonly used in the process, machine learning techniques

are increasingly useful in various fields in analyzing and evaluating phenomena that are difficult to explain due to the complexity of social and economic systems including the environment and the accumulation of vast amounts of data. However, there are still areas that require explanation and understanding in securing reliability for analysis and verifying the model. In particular, the reliability of climate change-related models as in this study is more important because uncertainty is always present. Therefore, in order to secure the reliability of techniques such as machine learning, efforts must be made to reduce uncertainty from data collection to analysis. Therefore, in this study, two processes were performed to secure model uncertainty and reliability: 1) the uncertainty of the model was considered by comparing various models, and 2) a myriad of running times to consider the number of different cases.

In addition, this study focuses on the relationship between climate change and disasters. Looking at the process of 'prevention-preparation-response-recovery' applied to South Korea's traditional disaster management, the scope of the study covers from ‘prevention’ to ‘response’. Although the first three steps are important, the recovery process is also very important, and the most necessary discussion in this process is related to risk financing and insurance. It is necessary for many international agreements and initiatives such as the Sendai Framework for Disaster Risk Reduction, the Sustainable Development Goals (SDGs), or the Paris Climate Agreement (Reguero et al., 2020).

Since damage caused by massive natural disasters, such as floods, is difficult to collateralize in private insurance, the development and sales of related insurance products have not been

active. if natural disasters occur more frequently in the future, resulting in a continuous increase in physical and human damage, the stability of the national economy and people's lives must be planned through insurance coverage. For instance, parametric insurance can be introduced, which pays out a flat insurance amount upon the occurrence of a certain trigger. Parametric insurance is a form of insurance-linked security, such as collateralized reinsurance or catastrophe bond, which was developed nearly 20 years ago to cover large risks, such as massive natural disasters, to which conventional insurance alone cannot respond. Therefore, if parametric insurance would be used properly, private insurance can help respond to risks that could not be covered in the past, including floods. Furthermore, from the standpoint of insurance subscribers, losses caused by floods can be compensated for to a certain extent. In this respect, product operation in terms of policies warrants examination in the future. However, because parametric insurances do not cover actual losses, the so-called protection gap cannot be avoided, and the subscribers cannot be fully compensated for such losses.

Bibliography

Chapter 2

Ashraful Islam, M., Mitra, D., Dewan, A., & Akhter, S. H. (2016). Coastal multi-hazard vulnerability assessment along the Ganges deltaic coast of Bangladesh-A geospatial approach. In Ocean and Coastal Management (pp. 127, 1-15).

Azam, M., Kim, H. S., & Maeng, S. J. (2017). Development of flood alert application in Mushim stream watershed Korea. International Journal of Disaster Risk Reduction, 21(July 2016), 11-26.

Balica, S. F., Wright, N. G., & van der Meulen, F. (2012). A flood vulnerability index for coastal cities and its use in assessing climate change impacts. In Natural Hazards (Vol. 64, Issue 1).

Berz, G., Kron, W., Loster, T., Rauch, E., Schimetschek, J., Schmieder, J., Siebert, A., Smolka, A., & Wirtz, A. (2001). World map of natural hazards – a global view of the distribution and intensity of significant exposures. Natural Hazards, 23(2-3), 443-465.

Bhable, S., Kayte, S., Mali, S., Kayte, J. N., & Maher, R. (2015). A Review Paper on Coastal Hazard. International Journal of Engineering Research and Applications, 5(12), 83-93.

Bhavsar, H., & Ganatra, A. (2012). A Comparative Study of Training Algorithms for Supervised Machine Learning. International Journal of Soft Computing and Engineering, 2(4), 74-81.

BREIMAN, L. (2001). Random Forests. Machine Learning, 45, 5-32.

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. Machine Leaming, 20, 273-297.

Danades, A., Pratama, D., Anggraini, D., & Anggriani, D. (2017). Comparison of accuracy level K-Nearest Neighbor algorithm and Support Vector Machine algorithm in classification water quality status. 2016 6th International Conference on System Engineering and Technology (ICSET), 137-141.

Dwarakish, G. S., Vinay, S. A., Natesan, U., Asano, T., Kakinuma, T., Venkataramana, K., Pai, B. J., & Babita, M. K. (2009). Coastal vulnerability assessment of the future sea level rise in Udupi coastal zone of Karnataka state, west coast of India. Ocean and Coastal Management, 52(9), 467-478.

Eilander, D., Couasnon, A., Ikeuchi, H., Muis, S., Yamazaki, D., Winsemius, H., & Ward, P. J. (2020). The effect of surge on riverine flood hazard and impact in deltas globally. Environmental Research Letters.

Giardino, A., Nederhoff, K., & Vousdoukas, M. (2018). Coastal hazard risk assessment for small islands: assessing the impact of climate change and disaster reduction measures on Ebeye (Marshall Islands). Regional Environmental Change, 18(8), 2237-2248.

Han, H. J., Kim, J. H., Chung, S. E., Park, J. H., & Cheong, H. K. (2018). Estimation of the national burden of disease and vulnerable population associated with natural disasters in Korea: Heavy precipitation and typhoon. Journal of Korean Medical Science, 33(49), 1-3.

Hao, T., Elith, J., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2019). A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Diversity and Distributions, 25(5), 839-852.

Harefa, J., & Pratiwi, M. (2016). Comparison Classifier: Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) In Digital Mammogram Images. Juisi, 02(02), 35-40.

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, 21(9), 1263-1284.

Hoch, J. M., Eilander, D., Ikeuchi, H., Baart, F., & Winsemius, H. C. (2019). Evaluating the impact of model complexity on flood wave propagation and inundation extent with a hydrologic-hydrodynamic model coupling framework. Natural Hazards and Earth System Sciences, 19(8), 1723-1735.

IPCC. (2007). Climate change, impacts, adaptation and vulnerability Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, ed M L Parry, O F Canziani, J P Palutikof, P J van der Linden and C E Hanson (New York: Cambridge University Press)

IPCC. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Jadhav, S. D., & Channe, H. P. (2016). Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques. International Journal of Science and Research (IJSR), 5(1), 1842-1845.

Kim, J., Kim, B.-S., & Savarese, S. (2012). Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines. Applied Mathematics in Electrical and Computer Engineering, 133-138.

Klein, R. J. T. and Nicholls, R. J. (1999). Assessment of coastal vulnerability to climate change Ambio 28 182-7

Klein, R. J. T., Nicholls, R. J., & Thomalla, F. (2003). Resilience to natural hazards: How useful is this concept? Environmental Hazards, 5(1), 35-45.

Kleint, R. J. T., Nicholls, R. J., Ragoonaden, S., Capobianco, M., Aston, J., & Buckley, E. N. (2001). Technological options for adaptation to climate change in coastal zones. Journal of Coastal Research, 17(3), 531-543.

Knutti, R., & Sedláček, J. (2013). Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change, 3(4), 369-373.

Korea Meteorological Administration. (2011). Korea Climate Change White Paper pp 1-117

Kourgialas, N. N., & Karatzas, G. P. (2011). Gestion des inondations et méthode de modélisation sous SIG pour évaluer les zones d'aléa inondation-une étude de cas. Hydrological Sciences Journal, 56(2), 212-225.

Kundzewicz, Z. W. et al (2014). Flood risk and climate change: global and regional perspectives. Hydrol. Sci. J., 59 1-28.

Lilai, X., Yuanrong, H., Wei, H., & Shenghui, C. (2016). A multi-dimensional integrated approach to assess flood risks on a coastal city, induced by sea-level rise and storm tides. Environmental Research Letters, 11(1).

L´opez-Serrano, P. M., L´opez-S´anchez, C. A., ´Alvarez-Gonz´alez, J. G., and García-Guti´errez, J. (2016). A comparison of ML techniques applied to Landsat-5 TM spectral data for biomass estimation. Can. J. Remote Sens., 42 690-705.

Mahendra, R. S., Mohanty, P. C., Srinivasa Kumar, T., Shenoi, S. S. C., & Nayak, S. R. (2010). Coastal Multi-Hazard Vulnerability Mapping: A Case Study Along The Coast of Nellore District, East Coast of India. Italian Journal of Remote Sensing.

Ministry of the Interior and Safety. (2016). Statistical Yearbook of Natural Disaster

Muis, S., Apecechea, M. I., Dullaart, J., de Lima Rego, J., Madsen, K. S., Su, J., Yan, K., & Verlaan, M. (2020). A High-Resolution Global Dataset of Extreme Sea Levels, Tides, and Storm Surges, Including Future Projections. Frontiers in Marine Science, 7(April), 1-15.

Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding - A global assessment. PLoS ONE, 10(3).

Nicholls, R. J., & Cazenave, A. (2010). Sea-level rise and its impact on coastal zones. Science, 328(5985), 1517-1520.

Oh, H. M., Jeong, K. Y., Kim, H. K., Lee, E., Hwang, S. M., Kim, S. M., & Kang, T. S. (2020). Wave Risk Assessment on Coastal Areas in Korea. In N. Trung Viet, D. Xiping, & T. Thanh Tung (Eds.), APAC 2019 (pp. 1351-1358). Springer Singapore.

Pantusa, D., D'Alessandro, F., Riefolo, L., Principato, F., & Tomasicchio, G. R. (2018). Application of a coastal vulnerability index. A case study along the Apulian Coastline, Italy. Water (Switzerland), 10(9), 1-16.

Parker, W. S. (2013). Ensemble modeling, uncertainty and robust predictions. Wiley Interdisciplinary Reviews: Climate Change, 4(3), 213-223.

Kinnerkar, K. P. R. (2016). A Comparative Study of Machine Learning Algorithms applied to Predictive Breast Cancer Data. International Journal of Science and Research (IJSR), 5(9), 1550-1553.

Rebentrost, P., Mohseni, M., & Lloyd, S. (2014). Quantum support vector machine for big data classification. Physical Review Letters, 113(3), 1-5.

Sahana, M., & Sajjad, H. (2019). Vulnerability to storm surge flood using remote sensing and GIS techniques: A study on Sundarban Biosphere Reserve, India. Remote Sensing Applications: Society and Environment, 13(November), 106-120.

Sankari, T. S., Chandramouli, A. R., Gokul, K., Surya, S. S. M., & Saravanavel, J. (2015). Coastal Vulnerability Mapping Using Geospatial Technologies in Cuddalore-Pichavaram Coastal Tract, Tamil Nadu, India. Aquatic Procedia, 4(Icwrcoe), 412-418.

Saxena, S., Geethalakshmi, V., & Lakshmanan, A. (2013). Development of habitation vulnerability assessment framework for coastal hazards: Cuddalore coast in Tamil Nadu, India-A case study. Weather and Climate Extremes, 2, 48-57.

Szlafsztein, C., & Sterr, H. (2007). A GIS-based vulnerability assessment of coastal natural hazards, state of Pará, Brazil. Journal of Coastal Conservation, 11(1), 53-66.

Thanh Noi, P., & Kappas, M. (2017). Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors (Basel, Switzerland), 18(1).

Simon Tong, & Koller, D. (2001). Support Vector Machine Active Learning with Applications to Text Classification. Journal of Machine Learning Research, 45-66.

Tran, P., Shaw, R., Chantry, G., & Norton, J. (2008). GIS and local knowledge in disaster management: a case study of flood risk mapping in thua thien hue province, Vietnam. Disasters, 33(1), 152-169.

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., & Feyen, L. (2018). Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. Nature Communications, 9(1), 1-12.

Vousdoukas, M. I., Voukouvalas, E., Mentaschi, L., Dottori, F., & Giardino, A. (2016). Developments in large-scale coastal flood hazard mapping. Natural Hazards and Earth System Sciences, 16, 1841-1853.

Wahl, T., N. G. Plant, and J. W. L. (2016). Probabilistic assessment of erosion and flooding risk in the northern Gulf of Mexico. Journal of Geophysical Research: Oceans, 121, 3029-3043.

Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., Veldkamp, T. I. E., Winsemius, H. C., & Wahl, T. (2018). Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. Environmental Research Letters, 13(8).

Wood, S. N. (2020). Inference and computation with generalized additive models and their extensions. In Test (Vol. 29, Issue 2). Springer Berlin Heidelberg.

Van Den Hurk, B., Van Meijgaard, E., De Valk, P., Van Heeringen, K. J., & Gooijer, J. (2015). Analysis of a compounding surge and precipitation event in the Netherlands. Environmental Research Letters, 10(3).

Yoon, D. K., Kang, J. E., & Brody, S. D. (2016). A measurement of community disaster resilience in Korea. Journal of Environmental Planning and Management, 59(3), 436-460.

Chapter 3

Abedini, M., Ghasemian, B., Shirzadi, A., Shahabi, H., Chapi, K., Pham, B. T., Bin Ahmad, B., & Tien Bui, D. (2019). A novel hybrid approach of Bayesian Logistic Regression and its ensembles for landslide susceptibility assessment. Geocarto International, 34(13), 1427-1457.

Akgun A. (2012). A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at Izmir, Turkey. Landslides. 9:93-106.

Barría, P., Cruzat, M. L., Cienfuegos, R., Gironás, J., Escauriaza, C., Bonilla, C., Moris, R., Ledezma, C., Guerra, M., Rodríguez, R., & Torres, A. (2019). From Multi-Risk Evaluation to Resilience Planning: The Case of Central Chilean Coastal Cities. Water, 11(3), 572.

Bhavsar, H., & Ganatra, A. (2012). A Comparative Study of Training Algorithms for Supervised Machine Learning. International Journal of Soft Computing and Engineering, 2(4), 74–81.

BREIMAN, L. (2001). Random Forests. Machine Learning, 45, 5–32.

Bui, D. T., Shirzadi, A., Shahabi, H., Geertsema, M., Omidvar, E., Clague, J. J., Pham, B. T., Dou, J., Asl, D. T., Ahmad, B. Bin, & Lee, S. (2019). New ensemble models for shallow landslide susceptibility modeling in a semi-aridwatershed. Forests, 10(9).

Chang, K. T., Merghadi, A., Yunus, A. P., Pham, B. T., & Dou, J. (2019). Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. Scientific Reports, 9(1), 1–21.

Chen, W., Shirzadi, A., Shahabi, H., Ahmad, B. Bin, Zhang, S., Hong, H., & Zhang, N. (2017a). A novel hybrid artificial intelligence approach based on the rotation forest ensemble and naïve Bayes tree classifiers for a landslide susceptibility assessment in Langao County, China. Geomatics, Natural Hazards and Risk, 8(2), 1955–1977.

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., Duan, Z., & Ma, J. (2017b). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena, 151, 147–160.

Chen, W., Pourghasemi, H. R., & Naghibi, S. A. (2018). Prioritization of landslide conditioning factors and its spatial modeling in Shangnan County, China using GIS-based data mining algorithms. Bulletin of Engineering Geology and the Environment, 77(2), 611–629.

Cracknell, M. J., & Reading, A. M. (2014). Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Computers and Geosciences, 63, 22–33.

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. Machine Leaming, 20, 273–297.

DeFries, R. S., & Chan, J. C. W. (2000). Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sensing of Environment, 74(3), 503–515.

Dou, J., Yamagishi, H., Pourghasemi, H. R., Yunus, A. P., Song, X., Xu, Y., & Zhu, Z. (2015). An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan. Natural Hazards, 78(3), 1749–1776.

Dou, J., Yunus, A. P., Tien Bui, D., Merghadi, A., Sahana, M., Zhu, Z., Chen, C. W., Khosravi, K., Yang, Y., & Pham, B. T. (2019a). Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. Science of the Total Environment, 662, 332–346.

Dou, J., Yunus, A. P., Tien Bui, D., Sahana, M., Chen, C.-W., Zhu, Z., Wang, W., & Pham, B. T. (2019b). Evaluating GIS-Based Multiple Statistical Models and Data Mining for Earthquake and Rainfall-Induced Landslide Susceptibility Using the LiDAR DEM. Remote Sensing, 11(6), 638.

Dou, J., Yunus, A. P., Bui, D. T., Merghadi, A., Sahana, M., Zhu, Z., Chen, C. W., Han, Z., & Pham, B. T. (2020a). Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides, 17(3), 641–658.

Dou, J., Yunus, A. P., Merghadi, A., Shirzadi, A., Nguyen, H., Hussain, Y., Avtar, R., Chen, Y., Pham, B. T., & Yamagishi, H. (2020b). Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. Science of the Total Environment, 720(February), 137320.

García-Rodríguez, M. J., Malpica, J. A., Benito, B., & Díaz, M. (2008). Susceptibility assessment of earthquake-triggered landslides in El Salvador using logistic regression. Geomorphology, 95(3–4), 172–191.

Goetz, J. N., Brenning, A., Petschko, H., & Leopold, P. (2015). Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers and Geosciences, 81, 1–11.

Gómez, H., & Kavzoglu, T. (2005). Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Engineering Geology, 78(1–2), 11–27.

IPCC. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Hao, T., Elith, J., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2019). A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Diversity and Distributions, 25(5), 839–852.

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, 21(9), 1263-1284.

Hong, H., Chen, W., Xu, C., Youssef, A. M., Pradhan, B., & Tien Bui, D. (2017a). Rainfall-induced landslide susceptibility assessment at the Chongren area (China) using frequency ratio, certainty factor, and index of entropy. Geocarto International, 32(2), 139-154.

Hong, H., Liu, J., Zhu, A. X., Shahabi, H., Pham, B. T., Chen, W., Pradhan, B., & Bui, D. T. (2017b). A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the Wuning area (China). Environmental Earth Sciences, 76(19), 1-19.

Korea Meteorological Administration official website: https://www.weather.go.kr

Jadhav, S. D., & Channe, H. P. (2016). Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques. International Journal of Science and Research (IJSR), 5(1), 1842-1845.

Kim, J., Kim, B.-S., & Savarese, S. (2012). Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines. Applied Mathematics in Electrical and Computer Engineering, 133-138.

Knutti, R., & Sedláček, J. (2013). Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change, 3(4), 369-373.

Kornejady, A., Ownegh, M., & Bahremand, A. (2017). Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. Catena, 152, 144-162.

Kumar, S., Snehmani, Srivastava, P. K., Gore, A., & Singh, M. K. (2016). Fuzzy-frequency ratio model for avalanche susceptibility mapping. International Journal of Digital Earth, 9(12), 1168-1184.

Lee, S., & Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. Landslides, 4(1), 33-41.

Malamud, B. D., Turcotte, D. L., Guzzetti, F., Reichenbach, P. (2004). Landslide inventories and their statistical properties. Earth Surf Process Landforms. 29:687-711.

Ministry of the Interior and Safety official website: https://www.mois.go.kr

Naghibi, S. A., Ahmadi, K., & Daneshi, A. (2017). Application of Support Vector Machine, Random Forest, and Genetic Algorithm Optimized Random Forest

Models in Groundwater Potential Mapping. Water Resources Management, 31(9), 2761-2775.

Nsengiyumva, J. B., Luo, G., Nahayo, L., Huang, X., & Cai, P. (2018). Landslide susceptibility assessment using spatial multi-criteria evaluation model in Rwanda. International Journal of Environmental Research and Public Health, 15(2).

O'Brien, R.M. (2007). A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673-690.

Park, S. J., & Lee, D. K. (2020). Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. Environmental Research Letters, 15(9).

Parker, W. S. (2013). Ensemble modeling, uncertainty and robust predictions. Wiley Interdisciplinary Reviews: Climate Change, 4(3), 213-223.

Pham, B. T., Bui, D., Prakash, I., & Dholakia, M. B. (2016). Evaluation of predictive ability of support vector machines and naive Bayes trees methods for spatial prediction of landslides in Uttarakhand state (India) using GIS. Journal of Geomatics, 10(1), 71-79.

Pham, B. T., Tien Bui, D., Prakash, I., & Dholakia, M. B. (2017). Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena, 149, 52-63.

Pham, B. T., & Prakash, I. (2019). A novel hybrid model of Bagging-based Naïve Bayes Trees for landslide susceptibility assessment. Bulletin of Engineering Geology and the Environment, 78(3), 1911-1925.

Pham, B. T., Prakash, I., Dou, J., Singh, S. K., Trinh, P. T., Tran, H. T., Le, T. M., Van Phong, T., Khoi, D. K., Shirzadi, A., & Bui, D. T. (2020). A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. Geocarto International, 35(12), 1267-1292.

Polykretis, C., Chalkias, C., & Ferentinou, M. (2019). Adaptive neuro-fuzzy inference system (ANFIS) modeling for landslide susceptibility assessment in a Mediterranean hilly area. Bulletin of Engineering Geology and the Environment, 78(2), 1173-1187.

Pourghasemi, H. R., Kariminejad, N., Amiri, M., Edalat, M., Zarafshar, M., Blaschke, T., & Cerda, A. (2020). Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. Scientific Reports, 10(1), 3203.

Rebentrost, P., Mohseni, M., & Lloyd, S. (2014). Quantum support vector machine for big data classification. Physical Review Letters, 113(3), 1–5.

Tien Bui, D., Pradhan, B., Lofman, O., Revhaug, I., & Dick, O. B. (2012). Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro−fuzzy inference system and GIS. Computers and Geosciences, 45, 199–211.

Tong, S. & Koller, D. (2001). Support Vector Machine Active Learning with Applications to Text Classification. Journal of Machine Learning Research, 45–66.

UNISDR. (2015). Making Development Sustainable: The Future of Disaster Risk Management. Global Assessment Report on Disaster Risk Reduction. Geneva, Switzerland: United Nations Office for Disaster Risk Reduction (UNISDR).

Wang, W., He, Z., Han, Z., Li, Y., Dou, J., & Huang, J. (2020). Mapping the susceptibility to landslides based on the deep belief network: a case study in Sichuan Province, China. Natural Hazards, 103(3), 3239–3261.

Yilmaz, I. (2009). Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat−Turkey). Computers and Geosciences, 35(6), 1125–1138.

Yilmaz, I. (2010). The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability and artificial neural networks. Environmental Earth Sciences, 60(3), 505–519.

Zare, M., Pourghasemi, H. R., Vafakhah, M., & Pradhan, B. (2013). Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: A comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. Arabian Journal of Geosciences, 6(8), 2873–2888.

Zhou, C., Yin, K., Cao, Y., Ahmed, B., Li, Y., Catani, F., & Pourghasemi, H. R. (2018). Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. Computers and Geosciences, 112(November 2017), 23–37.

Chapter 4

Ahiablame, L., & Shakya, R. (2016). Modeling fl ood reduction effects of low impact development at a watershed scale. Journal of Environmental Management, 171, 81–91.

Arjenaki, M. O., Sanayei, H. R. Z., Heidarzadeh, H., & Mahabadi, N. A. (2021). Modeling and investigating the effect of the LID methods on collection network of urban runoff using the SWMM model (case study: Shahrekord City). In Modeling Earth Systems and Environment (Vol. 7, Issue 1). Springer Science and Business Media Deutschland GmbH.

Barnard, P. L. et al. (2015). Coastal vulnerability across the Pacific dominated by El Niño/Southern Oscillation. Nature Geoscience. 8(10), 801–807.

Bhavsar, H., & Ganatra, A. (2012). A Comparative Study of Training Algorithms for Supervised Machine Learning. International Journal of Soft Computing and Engineering, 2(4), 74–81.

BREIMAN, L. (2001). Random Forests. Machine Learning, 45, 5–32.

Caldwell, P. C., Vitousek, S. & Aucan, J. P. (2009). Frequency and duration of coinciding high surf and tides along the North Shore of Oahu, Hawaii, 1981–2007. Journal of Coastal Research. 734–743.

Cheong, SM., Silliman, B., Wong, P. et al. (2013). Coastal adaptation with ecological engineering. Nature Climate Change 3, 787–791.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., Payne, A. J., Pfeffer, W. T., Stammer, D., Unnikrishnan A. S. (2013). Sea level change, in Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, T. F. Stocker et al., Eds. (Cambridge Univ. Press, 2013), pp. 1137–1216.

Creach, A., Bastidas-Arteaga, E., Pardo, S., & Mercier, D. (2020). Vulnerability and costs of adaptation strategies for housing subjected to flood risks: Application to La Guérinière France. Marine Policy, 117, 103438.

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification IEEE Trans. Inf. Theory 13 21–7

Déqué, M. (2007). Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. Global and Planetary Change, 57(1–2), 16–26.

Dong, S., Abolfathi, S., Salauddin, M., Tan, Z.H., Pearson, J.M. (2020). Enhancing climate resilience of vertical seawall with retrofitting – A physical modelling study. Applied Ocean Research, 103, 102331.

Duvat, V. (2013). Coastal protection structures in Tarawa atoll, Republic of Kiribati. Sustainability science, 8(3), 363–379.

Eilander, D., Couasnon, A., Ikeuchi, H., Muis, S., Yamazaki, D., Winsemius, H., & Ward, P. J. (2020). The effect of surge on riverine flood hazard and impact in deltas globally. Environmental Research Letters.

Ferreira, Ó., Plomaritis, T. A., & Costas, S. (2019). Effectiveness assessment of risk reduction measures at coastal areas using a decision support system: Findings from Emma storm. Science of the Total Environment, 657, 124-135.

IPCC. (2007). Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Eds., Cambridge University Press, Cambridge, UK, 976pp

IPCC. (2014). Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Barros, V.R., C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L.White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 688.

IPCC. (2018). Summary for Policymakers. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. World Meteorological Organization, Geneva, Switzerland, 32 pp.

IPCC. (2019). Special Report on the Ocean and Cryosphere in a Changing Climate (Geneva: IPCC)

IPCC. (2022). Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. In Press.

IUCN. (2012). The IUCN Programme 2013-2016. IUCN, Gland, p. 3

Jeong, D.Y., Kim, M., Song, K.H., and Lee, J.A. (2021). Planning a Green Infrastructure Network to Integrate Potential Evacuation Routes and the Urban Green Space in a Coastal City: The Case Study of Haeundae District, Busan, South Korea. Science of The Total Environment, 761, 143179

Jongman, B. (2018). Effective adaptation to rising flood risk. Nature Communications, 9(1), 9-11.

Kim, J., Kim, B. S., and Savarese, S. (2012). Comparing image classification methods: K-Nearest-Neighbor and support-vector-machines Applied Mathematics in Electrical and Computer Engineering pp 133-8

Kron, W. (2013). Coasts: The high-risk areas of the world. Natural Hazards, 66(3), 1363-1382.

Korea Meteorological Administration. (2020). Korea Climate Change Assessment Report 2020

Kulp, S. A., & Strauss, B. H. (2019). New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. Nature Communications, 10(1).

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C., Bates, P., ... & Macdonald, E. (2021). Causes, impacts and patterns of disastrous river floods. Nature Reviews Earth & Environment, 2(9), 592-609.

Ministry of Public Safety and Security. (2015). STATISTICAL YEARBOOK OF NATURAL DISASTER.

FEMA, (2018). Natural Hazard Mitigation Saves: Utilities and Transportation Infrastructure.

Li, F., Chen, J., Engel, B. A., Liu, Y., Wang, S., & Sun, H. (2021). Assessing the effectiveness and cost efficiency of green infrastructure practices on surface runoff reduction at an urban watershed in China. Water (Switzerland), 13(1).

National Institute of Meteorological Science. (2018). 100 Years of Climate Change on the Korean Peninsula.

Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding - A global assessment. PLoS ONE, 10(3).

Oh, H. M., Jeong, K. Y., Kim, H. K., Lee, E., Hwang, S. M., Kim, S. M., & Kang, T. S. (2020). Wave Risk Assessment on Coastal Areas in Korea. In N. Trung Viet, D. Xiping, & T. Thanh Tung (Eds.), APAC 2019 (pp. 1351-1358). Springer Singapore.

Park, S. J., & Lee, D. K. (2020). Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. Environmental Research Letters, 15(9).

Parker, W. S. (2013). Ensemble modeling, uncertainty and robust predictions. Wiley Interdisciplinary Reviews: Climate Change, 4(3), 213–223.

Potdar, K., & Kinnerkar, R. (2016). A Comparative Study of Machine Learning Algorithms applied to Predictive Breast Cancer Data. 5(9), 2013–2016.

Reguero, B. G., Losada, I. J., Díaz-Simal, P., Méndez, F. J., & Beck, M. W. (2015). Effects of climate change on exposure to coastal flooding in Latin America and the Caribbean. PLoS ONE, 10(7), 1–19.

Reguero, B. G., Beck, M. W., Bresch, D. N., Calil, J., & Meliane, I. (2018). Comparing the cost effectiveness of nature-based and coastal adaptation: A case study from the Gulf Coast of the United States. PLoS ONE, 13(4), 1–24.

Reguero, B. G., Beck, M. W., Schmid, D., Stadtmüller, D., Raepple, J., Schüssele, S., & Pfliegner, K. (2020). Financing coastal resilience by combining nature-based risk reduction with insurance. Ecological Economics, 169(November 2019).

Singhvi, A., Luijendijk, A. P., & van Oudenhoven, A. P. E. (2022). The grey – green spectrum: A review of coastal protection interventions. Journal of Environmental Management, 311(March), 114824.

Spalding, M. D., Ruffo, S., Lacambra, C., Meliane, I., Hale, L.Z., Shepard, C.C., Beck, M.W. (2014). The role of ecosystems in coastal protection: Adapting to climate change and coastal hazards. Ocean Coast. Manag. 90, 50–57

Tiggeloven, T., De Moel, H., Winsemius, H. C., Eilander, D., Erkens, G., Gebremedhin, E., ... & Ward, P. J. (2020). Global-scale benefit-cost analysis of coastal flood adaptation to different flood risk drivers using structural measures. Natural Hazards and Earth System Sciences, 20(4), 1025-1044.

Vitousek, S., Barnard, P. L., Fletcher, C. H., Frazer, N., Erikson, L., & Storlazzi, C. D. (2017). Doubling of coastal flooding frequency within decades due to sea-level rise. Scientific Reports, 7(1), 1–9.

Vousdoukas, M. I., Mentaschi, L., Hinkel, J., Ward, P. J., Mongelli, I., Ciscar, J. C., & Feyen, L. (2020). Economic motivation for raising coastal flood defenses in Europe. Nature communications, 11(1), 1-11.

Watson, C. S., White, N. J., Church, J. A., King, M. A., Burgette, R. J., & Legresy, B. (2015). Unabated global mean sea-level rise over the satellite altimeter era. Nature Climate Change, 5(6), 565–568.

World Bank. (2018). NATURE-BASED SOLUTIONS FOR DISASTER RISK MANAGEMENT.

Yang, K. (2008). Artificial Neural Networks(ANNs): A New Paradigm for Thermal Science and Engineering. Journal of Heat Transfer, 130(September), 1-19.

Yi, S., Sun, W., Heki, K. & Qian, A. (2015). An increase in the rate of global mean sea level rise since 2010. Geophysical Research Letters, 42(10), 3998-4006.

Zhang, B., Xie, G., Zhang, C., & Zhang, J. (2012). The economic benefits of rainwater-runoff reduction by urban green spaces: A case study in Beijing, China. Journal of Environmental Management, 100, 65.

Abstract in Korean

머신러닝 기법을 활용한 기후변화 영향에 따른 재해 리스크 평가

박 상 진
서울대학교 환경대학원 협동과정조경학 및
대학원 융합전공 스마트시티 글로벌 융합
논문지도교수: 이 동 근

기후 변화는 우리 세대에게 시급한 위협이다. 자연 재해는 기후 변화로 인해 더 잦은 빈도와 강력하게 발생하고 있어 예측불가성이 커져가고 있다. 특히, 한국의 자연재해는 대부분 기상 현상으로 인해 발생하는데, 지난 10 년간 재해로 인한 전체 피해는 주로 태풍(49%)과 호우(40%)에 기인하였다. 따라서 장기적으로 대비하기 위해서는 홍수, 산사태 등 호우와 관련된 위험을 분석하고 평가하는 위험관리가 필요하다.

따라서 본 논문의 주요 연구질문은 다음과 같다: 1) 기후변화로 인한 복잡한 상황에서 다양한 요인을 고려하여 미래의 잠재적 위험을 어떻게 예측할 것인가, 2) 이러한 위험을 줄이기 위해 어떤 노력을 하는 것이 지속가능한가?. 먼저 연안 홍수, 산사태 등 복합적 영향의 미래 위험도를 평가하기 위해 첫째, 최근 연구에서 널리 활용되고 있는 다중 머신러닝(ML) 알고리즘을 확률론적 접근 방식으로 활용하여 현재의 위험도를 분석하였다. 다양한 RCP 기후변화 시나리오 및 지역 기후 모델에 따른 예측 강우량을 고려하여 미래 위험을 추정했습니다. 둘째, 기후변화 영향으로 인한 재난위험 대응을 위한 적응전략의 실효성을 평가하기 위하여, 적응전략으로 중요한 역할을 하는 녹지, 방파제 등

구조적 대책의 효과성과 지속가능성을 여러 적응경로로 나눠 연안침수에 대한 위험저감을 평가하였다.

연구의 결과는 미래의 위험 지역을 식별하고 위험 관리를 위한 의사 결정 과정, 그리고 토지 이용 계획 및 의사 결정 프로세스를 포함한 재난 감소 및 관리 조치에 대해 지원 가능할 것이다.

Keyword: 기후변화 영향, 재해 리스크, 리스크 관리, 기후변화 시나리오, 해수면 상승, 연안통합관리, 산사태 민감도, 자연기반 솔루션, 적응전략

Student Number: 2019-39641