Forest fire susceptibility assessment using google earth engine in Gangwon-do, Republic of Korea

Yong Piao\textsuperscript{a}, Dongkun Lee\textsuperscript{b,c}, Sangjin Park\textsuperscript{d}, Ho Gul Kim\textsuperscript{e} and Yihua Jin\textsuperscript{f}

\textsuperscript{a}Graduate School of Environmental Studies, Seoul National University, Gwanak-ro, Gwanak-gu, Seoul, Korea; \textsuperscript{b}Research Institute of Agriculture Life Science, Seoul National University, Gwanak-ro, Gwanak-gu, Seoul, Korea; \textsuperscript{c}Department of Landscape Architecture and Rural System Engineering, Seoul National University, Seoul, Korea; \textsuperscript{d}Interdisciplinary Program in Landscape Architecture & Transdisciplinary Program in Smart City Global Convergence, Seoul National University, Seoul, Korea; \textsuperscript{e}Department of Human Environment Design, Cheongju University, Cheongju, ROK; \textsuperscript{f}Department of Landscape Architecture, Agricultural college, Yanbian University, Yanji, China

ABSTRACT
Forest fires are one of the most frequently occurring natural hazards, causing substantial economic loss and destruction of forest cover. As the Gangwon-do region in Korea has abundant forest resources and ecological diversity as Korea’s largest forest area, spatial data on forest fire susceptibility of the region are urgently required. In this study, a forest fire susceptibility map (FFSM) of Gangwon-do was constructed using Google Earth Engine (GEE) and three machine learning algorithms: Classification and Regression Trees (CART), Random Forest (RF), and Boosted Regression Trees (BRT). The factors related to climate, topography, hydrology, and human activity were constructed. To verify the accuracy, the area under the receiver operating characteristic curve (AUC) was used. The AUC values were 0.846 (BRT), 0.835 (RF), 0.751 (CART). Factor importance analysis was performed to identify the important factors of the occurrence of forest fires in Gangwon-do. The results show that the most important factor in the Gangwon-do region is slope. A slope of approximately 17\degree (moderately steep) has a considerable impact on the occurrence of forest fires. Human activity and interference are the other important factors that affect forest fires. The established FFSM can support future efforts on forest resource protection and environmental management planning in Gangwon-do.

ARTICLE HISTORY
Received 26 October 2021
Accepted 13 January 2022

KEYWORDS
Natural hazard; Wildfire; Google earth engine; remote sensing; machine learning

1. Introduction
Forests cover about one-third of the Earth’s surface (Keenan et al. 2015) and are important environmental resources, producing approximately two-thirds of the world’s oxygen (McKinley et al. 2011). However, forests are sensitive to the impacts...
of climate change, human activity (McKinley et al. 2011), and various disasters such as forest fires, landslides, and pests (Lecina-Diaz et al. 2021). Forest fires cause severe damage to forests and surrounding communities because they spread rapidly to very large areas (Mollicone et al. 2006).

Forest fires are mainly caused by two factors: human activity (e.g., deforestation) (Cochrane 2001, Skole and Tucker 1993), and the natural environment (e.g., lightning strikes) (Tutin et al. 1996). Forest fires destroy forest ecosystems and cause massive economic loss. In addition, burned forests become severely damaged, and the large amounts of combustible wood, greatly increase the risk of recurrence of forest fires (Siegert et al. 2001). Therefore, it is crucial to take measures to monitor and prevent forest fires.

Forest fire susceptibility mapping is an important step in the prevention of damage caused by forest fire. Forest fire susceptibility maps (FFSMs) can help identify areas at high risk of forest fires (Jaiswal et al. 2002). In addition, the destruction of forest ecosystems, casualties, and economic damage can be mitigated by providing spatial information on the susceptibility to forest fires to managers and planners through FFSMs (Jaiswal et al. 2002, Pourghasemi et al. 2020). Therefore, FFSMs play an important role in managing the forest fire risk.

Several tools have been used for forest fire susceptibility analysis and mapping, including satellite imaging technology (Hernandez-Leal et al. 2006, Prosper-Laget et al. 1995), geographic information systems technology (Jaiswal et al. 2002, Teodoro and Duarte 2013), fire area simulators with probabilistic models (Krasnow et al. 2009), fuel moisture content (Chuvieco et al. 2004), multivariate logistic regression (Pourghasemi et al. 2016), and generalized additive models (Pourtaghi et al. 2016). In addition, multi-criteria decision-making has been applied to the analytic hierarchy process (Pourghasemi et al. 2016), and step-wise weight assessment ratio analysis has been performed (Pourghasemi et al. 2019).

Recent studies have reported the use of artificial neural networks and machine learning as modeling approaches to achieve high accuracy in forest fire susceptibility analysis: fuzzy logic (Abedi Gheshlaghi et al. 2020, Pourghasemi et al. 2016), artificial neural networks (Satir et al. 2016), adaptive neuro-fuzzy inference systems and ensemble models of metaheuristics algorithms (Moayedi et al. 2020, Pourghasemi et al. 2019), classification and regression trees (CART) (Amatulli et al. 2006), boosted regression trees (BRTs) (Pourghasemi et al. 2020, Pourtaghi et al. 2016), and random forest (RF) algorithms (Milanović et al. 2020, Pourtaghi et al. 2016). Machine learning algorithms can be used as soft computing models when statistical limitations are low (Pourghasemi et al. 2020).

Google Earth Engine (GEE, https://earthengine.google.com) is a cloud-based satellite image processing platform. GEE has high computational speed and can be easily applied to a broad area (Gorelick et al. 2017). GEE can be conveniently used in spatial data analysis because it is possible to efficiently process spatial data using codes and APIs (Liu et al. 2018, Sidhu et al. 2018). In addition, GEE supports data upload services so that users can freely process data through various functions and codes on the platform. The number of applications of GEE has been continuously increasing since 2013, and successful analyses have been conducted in various research fields,
such as land cover classification, forest and vegetation, ecosystems, and agriculture (Hu and Hu 2019, Oliphant et al. 2019, Piao et al. 2021, Tamiminia et al. 2020).

In Gangwon-do, South Korea, mountains cover approximately 81% of the total area. The topography is complex and curved, and forest resources are abundant. Many wide-ranging forest fires occur annually in Gangwon-do (Lee and Lee 2006, Shin and Lee 2004). Forest fires in this area have caused significant damage, including massive economic loss, casualties, and destruction of ecological cover. FFSMs can identify the occurrence of forest fires in the region and provide guidance for forest fire prevention and planning. Therefore, there is an urgent need for forest fire susceptibility maps in Gangwon-do, and to the best of our knowledge, there is very few research on FFSMs in Gangwon-do.

In this study, GEE and three machine learning algorithms (CART, RF, and BRT) were used to construct and evaluate an FFSM in Gangwon-do, South Korea. The occurrence of forest fires were assessed through the analysis of climate, topography, hydrology, human activity, and interference data in the area. The important factors in the Gangwon-do region and the influence of human interference and activity were ranked through spatial analysis. The results are expected to be used in the protection of forest resources, biodiversity conservation, and environmental management planning.

2. Materials and methods

2.1. Study site

As indicated in Figure 1 this study was conducted in Gangwon-do, which is a municipality located in the northeastern part of South Korea, bordering the East Sea to the east, Gyeonggi-do to the west, Chungcheongbuk-do, Gyeongsangbuk-do to the south, and North Korea to the north (DMZ line). Gangwon-do has the largest forest area in Korea, accounting for approximately 21% of Korea's forest areas. Approximately 81% of Gangwon-do is composed of forests, which are vulnerable to frequent forest fires.

Figure 1. South Korea, Gangwon-do study area and forest fire inventory map on forest cover map (Ministry of Environment, http://me.go.kr/).
Of the 10 catastrophic forest fires in Korea, 7 occurred in Gangwon-do, causing enormous damage and economic loss (Korea Forest Service 2004, https://www.forest.go.kr).

### 2.2. Materials

#### 2.2.1. Inventory of forest fires

It is important to understand and map the spatial distribution of natural disasters to study the spatial relationship between the location of occurrence of disasters and their causes (Pourghasemi et al. 2019). Area and location data are important for assessing and predicting susceptibility (Moayedi et al. 2020). Forest fire data were provided by the Korea Forest Service (https://www.forest.go.kr), and a total of 270 forest fire occurrence points were established. Of these 270 points, 70% were separated for training the model for forest fire susceptibility mapping, and 30% for testing the model. In addition, 270 non-fire area sample points were established and separated (Moayedi et al. 2020). A point where a forest fire occurred was designated as 1, and a point where a forest fire did not occur was designated as 0. The sample forest fire occurrence points are the ignition points, and the causes have been mostly identified as artificial ignition (Korea Forest Service 2004, https://www.forest.go.kr).

#### 2.2.2. Factors of forest fires

Disasters are caused by the interactions between various triggers (Van Westen et al. 2003, Verde and Zézere 2010). Based on previous studies (Moayedi et al. 2020, Pourghasemi et al. 2016, Pourghasemi et al. 2020) and the characteristics of the study region, elevation (m), slope (°), aspect, distance to roads (m), distance to rivers (m), distance to urban areas (m), rainfall amount (mm), annual average temperature (°C), drain density (m⁻¹), normalized difference vegetation index (NDVI), and topographic wetness index were used for forest fire susceptibility mapping. Table 1 shows a detailed description of the data sources of 11 influencing factors. Because the original data consisted of different points, lines, and polygons, all data used for forest fire

| Class          | Data Type                           | Source                                      | Resolution |
|----------------|-------------------------------------|---------------------------------------------|------------|
| Satellite Image| Landsat OLI                         | https://code.earthengine.google.com/        | 30 m       |
|                | Shuttle Radar Topography Mission V3 | https://code.earthengine.google.com/        | 30 m       |
| Geological data| River line map                      | Korea Ministry of Environment (http://water.nier.go.kr/) | Line data |
|                | Road line map                       | Korea National Spatial Data Infrastructure Portal (http://www.nnsi.go.kr/) | Polygon data |
|                | Rainfall and Temperature dataset    | Korea Meteorological Administration (https://data.kma.go.kr/) | Point data |
|                | Forest Fire list                    | Korea Forest Service (https://www.forest.go.kr/) | Point data |
|                | Land Use Land Cover map             | Korea Ministry of Environment (https://egis.me.go.kr/) | Polygon data |
|                | Urban Land Cover                    |                                             |            |
| Satellite Index| Normalized difference vegetation index (NDVI) | https://code.earthengine.google.com/ | 30 m       |
analysis were converted to a 30-m grid. Figure 2 shows the geographic and spatial information on 11 influencing factors used in the FFSM. The 11 factors were selected considering the characteristics of the Gangwon-do region, such as climate, topography, hydrology, human activity, and interference. Over the past 40 years, frequent fires have occurred in the urban and residential areas along major roads in
Gangwon-do (Bae 2018). Therefore, the characteristics of human activity and interference were reflected by linking the distance to roads (m) and distance to urban areas (m) with the forest fire occurrence points. Furthermore, it should be noted that when we constructed NDVI, in order to separate the forest cover from other covers in the study area, we used the NDVI value of the study area from April to July, and the temporal aggregation method to reduce cloud interference.

2.3. Methodology

2.3.1. Flow of study
The method used in this study includes three steps (Figure 3). In Step 1, pre-processing for attribute designation and coordinate system re-projection were carried out for the factors and forest fire samples built on the GEE platform. In Step 2, the FFSM was constructed using three machine learning algorithms (CART, RF, and BRT), and the accuracy of the prediction results was validated using the area under the receiver operating characteristic (ROC) curve (AUC). In step 3, susceptibility assessment was performed using the FFSM built using the most accurate model among the three machine learning algorithms.

2.3.2. Google earth engine platform
Google Earth Engine is a cloud-based platform that enables the rapid analysis and operation of spatial data (Gorelick et al. 2017). In this study, the susceptibility assessment of forest fires was conducted using the GEE platform. Although it is possible to import a large amount of spatial data in the GEE platform, additional factors should
be uploaded because there are no data from local governments of many countries. The method used in this study is described as follows:

1. We proceeded with a clip, reprojected and re-specified the scale for 11 factors uploaded to the GEE platform according to the characteristics of the study area.
2. All factors were superimposed and combined into one piece of data, and then the forest fire inventory map and attributes were connected.
3. The forest fire inventory map was divided into training (70%) and test (30%) datasets. The model was trained on the training dataset 500 times using the three machine learning models, and forest fire susceptibility was predicted.
4. The test set and ROC curve were used to verify the accuracy of the FFSM.
5. After validation, the trained model was used for the factor importance analysis.
6. The assessment of forest fires in the study area was carried out through the spatial distribution and location of the results of each model.

2.3.3. Machine learning algorithms

The FFSM was conducted using three machine learning algorithms: CART, RF, and BRT (Naghibi et al. 2016, Pourghasemi and Rahmati 2018, Shabani et al. 2021). We compared these algorithms after predicting the FFSM results. All three classifier models were built on the GEE platform.

The CART algorithm proposed by Breiman et al. (Breiman et al. 1984) uses a non-parametric nonlinear technique called binary partitioning algorithm as decision tree-based classifiers. In this binary partitioning algorithm, CART recursively splits the data to predict the distribution of the main target properties and compare and divide them between two child nodes (Breiman et al. 1984).

The RF algorithm proposed by Breiman (Breiman 2001) is currently one of the most widely used machine learning models. The RF algorithm is an ensemble learning method used for classification and regression analysis (Naghibi et al. 2016). RF includes bootstrap aggregation (bagging) and random feature selection. In RF, multiple samples extracted through the bagging of training samples are combined with a classification tree, and prediction is performed through a voting process (Breiman 2001). Compared with CART, RF resamples data through replacement to increase the diversity between classification trees (Shabani et al. 2021).

BRT (or stochastic gradient boosting) proposed by Friedman (Friedman 2001) makes predictions through a combination of statistical regression trees and machine learning techniques. BRT fits complex nonlinear relationships using a coupling method; i.e., BRT fits multiple simple models and combines them for spatial prediction. Details of BRT are described in Elith et al. (Elith et al. 2008).

2.3.4. Validation and comparison

In this study, the AUC was used to verify the accuracy of the FFSMs generated using the CART, RF, and BRT models. The AUC is a widely used method for evaluating the accuracy and performance of prediction models (Naghibi et al. 2016, Park and Lee 2020, Pourghasemi et al. 2020, Shabani et al. 2021) by balancing two ratios: a false-positive rate (specificity) on the X-axis and a true positive rate (sensitivity) on
the Y-axis (Fawcett 2006). The AUC value ranges from 0.5 to 1.0; when the model does not predict well, the AUC value is close to 0.5, and when the model predicts well, it is close to 1. In general, AUC values greater than 0.8 indicate high prediction performance (Nachappa et al. 2020).

2.3.5. Assessment importance of factor
The importance of the selected 11 factors can be used to identify the factors that have the most influence on the occurrence of forest fires in the study area and to help decision makers make plans for preventing forest fires (Pourghasemi et al. 2020). We confirmed the priority of the factors by exploring the importance of factors for each machine learning model. After extracting the trained model information from the GEE platform’s `.explain()` code, the importance information was evaluated through the `ee.Dictionary().get('importance')` code. The importance rankings of the three models were compared on the GEE platform using the extracted importance information.

3. Results
3.1. Validation and comparison of FFSMs
Figure 4 shows the accuracy assessment results of each FFSM. For the CART model, the AUC was 0.751 (accuracy = 75.1%); for the RF model, the AUC was 0.835 (accuracy = 83.5%); for the BRT model, the AUC was 0.846 (accuracy = 84.6%). Among the three models, RF and BRT had an AUC value ≥0.8, which indicates high prediction accuracy (Nachappa et al. 2020).
3.2. Forest fire susceptibility mapping

The forest fire susceptibility maps (FFSM) predicted using the three models were exported from the GEE platform and then regenerated in the GIS environment. Susceptibility probability values range from 0 to 1; the closer they are to 1, the more susceptible the space is. In this study, each FFSM was classified into regions with five different susceptibility classes (Very Low, Low, Moderate, High, and Very High) using the natural break classification method (Liu and Duan 2018, Moayedi et al. 2020, Pourghasemi et al. 2019).

Figure 5 shows the FFSMs of the CART, RF, and BRT models. From a spatial point of view, most areas susceptible to forest fires in Gangwon-do are clustered. This can be attributed to the urban areas and croplands where human activities are actively located at the center (Cochrane 2001, Siegert et al. 2001). The northern and southern regions of Gangwon-do have very low susceptibility to forest fires because of the mountains and limited agricultural and human activity (Cochrane 2003). RF and BRT show similar FFSM results, but BRT predictions are more concentrated in areas susceptible to forest fires than RF. Therefore, the FFSM created using BRT has

Figure 5. Forest fire susceptibility maps (FFSM) using (a) Classification and Regression Trees (CART), (b) Random Forest (RF), and (c) Boosted Regression Trees (BRT) algorithms.
more non-susceptible areas (Very Low) than RF, and a more concentrated distribution of susceptible areas (Very High) can provide more detailed information for disaster management and assessment.

Table 2 shows the area and ratio of the susceptibility classes for each machine learning algorithm. In the study area, 43.33% (CART), 15.55% (RF), and 23.11% (BRT) of the Very High susceptibility class were occupied. Although RF and BRT show similar occupancy rates in High and Low classes, BRT is more concentrated than RF in areas least susceptible to forest fires (Very Low) and areas most susceptible to forest fires (Very High). This is the same as the spatially confirmed result shown in Figure 5.

### 3.3. Important factors of FFSMs

It is important to understand the relationship between the FFSM and each factor and to evaluate the importance of factors in understanding the occurrence of forest fires in the study area (Pourghasemi et al. 2020). Figure 6 shows the importance of 11 selected factors for the FFSMs of CART, RF, and BRT. In CART, the importance of aspect is significantly higher than the other factors. In RF, slope has the highest importance, topographic wetness index has the second, and drain density has the third priority. In BRT, all factors impact the FFSM. Among them, slope has the highest importance, followed by distance to rivers and drain density. In Gangwon-do, both RF and BRT slopes have the greatest influence on forest fires because the spread rate of forest fires increases as the slope of the forest fire increases (Kushla and Ripple 1997, Lentile et al. 2006). In addition, the two factors related to human activity and interference in BRT, the distance to the roads (m) ranks the fourth in importance, and the distance to urban areas (m) ranks the tenth. In Gangwon-do, the distance to roads is an important factor, and one of the main causes of forest fires is vehicle fires on roads in steep slopes (Bae 2018). In view of these results, in predicting and planning for forest fires in Gangwon-do, establishing regulations and protection plans for roads near winding mountain areas plays an important role in reducing the occurrence of forest fires.

### 4. Discussion

#### 4.1. Advantages of spatial analysis on google earth engine

In this study, 11 factors related to forest fires were established according to the characteristics of the study site in Gangwon-do, Korea. Then, forest fire susceptibility...
mapping was carried out through the data upload service and coding system of the GEE platform. We also performed an accuracy assessment and importance analysis of the results using various built-in functions and algorithms of the GEE. In the model developed using GEE, the time used for analysis is in seconds (s), and the analysis results can be directly downloaded through a connection to Google Drive (https://drive.google.com/). Currently, GEE is focused more on remote sensing and the pre-processing of large volumetric satellite image data, land cover classification, and forest/vegetation/agriculture data (Tamiminia et al. 2020). In this study, susceptibility analysis of forest fires was completed on the GEE platform using spatial data and machine learning algorithms. This shows that GEE can be independent of time and computer specifications while enabling hazard susceptibility mapping with excellent accuracy. This spatial analysis capability of GEE can be utilized in other studies in addition to hazard susceptibility analysis to be able to conduct research that requires certain time and computer specifications, such as species distribution model maps and potential vegetation types.

4.2. Impact of human activities on Forest fires and on biodiversity

Forest fires are natural hazards most affected by anthropogenic interference, and approximately 80% of forest fires are caused by human interference (Cochrane 2001, McKinley et al. 2011, Mollicone et al. 2006, Skole and Tucker 1993). Therefore, the regulation and management of human activities in forest areas are important for reducing the incidence of and damage caused by forest fires. The FFSM results of this study confirmed that susceptibility of urban, cropland, and adjacent forest areas...
to forest fires is remarkably high. Conversely, susceptibility is low in areas with low human activity; i.e., areas far from urban areas, cropland, and deep forest, and also in forest protection areas such as Baekdu-daegan and national parks in the Gangwon-do area (Figure 7) because of the restricted human access to these areas. The FFSM shows that the occurrence of forest fires in the Baekdu-daegan Mountain Range and in the national park has remarkably low susceptibility.

The northern region of Gangwon-do, which is our study site, is adjacent to the DMZ, which is a special political region between South Korea and North Korea (Eun-Jin and Inae 2018). In areas adjacent to the DMZ, human activity is low owing to the treaties and agreements. In the FFSM (Figure 8), susceptibility to forest fires in the DMZ and adjacent areas was also remarkably low. This is consistent with the results of previous studies showing that forest fires are greatly affected by anthropogenic interference (Mollicone et al. 2006) and shows the accuracy and reliability of the FFSM of this study. These results imply the significance of taking forest fire prevention measures such as forest fire hazard education, hiker guidance, and an intensive management system for high-frequency visit periods (Shin and Lee 2004).

The knowledge of the spatial distribution of biodiversity is crucial for forest fire prevention and protection planning. In Gangwon-do, mountainous areas dominate the complex topography, which has abundant forests and high biodiversity (Shin and
Lee 2004). The ecological and natural map prepared by the Korean Ministry of Environment (Figure 9) shows important ecological values, naturalness, and landscape values of the natural environment of Gangwon-do (egis.me.go.kr). The first-grade zone in the ecological and natural map demonstrates the high biodiversity and the highest conservation value of the region. Compared to the FFSM of this study, although most first-grade areas are in Very Low and Low susceptibility areas, some of them are located in areas susceptible to forest fires. In Gangwon-do, there are designated forest protection areas (Figure 7), but further forest fire prevention and protection plans are required in relation to biodiversity. In addition to the economic damage, forest fires cause considerable destruction of the ecotone of forests and damage to biodiversity, which emphasizes the importance of the management and planning of protected areas against forest fires (Shin and Lee 2004). This also calls for nature-based solutions for ecosystem protection and sustainable management & restoration (Woo and Han 2020).

Figure 8. Low susceptibility to forest fires in the adjacent areas of the Korean Demilitarized Zone (DMZ). Adjacent areas of the DMZ: low human activity, less human settlement, fewer main roads.

Figure 9. Spatial distribution of the forest fire susceptibility map (FFSM) and the ecological map. Left: ecology map from Korean Ministry of Environment; Right: Overlay of the FFSM and the ecological map.
4.3. Spatial analysis of important factors of FFSMs

To understand the susceptibility to forest fires, it is important to perform spatial analysis of the factors of FFSMs. We spatially analyzed the slope, which is the most important factor in Gangwon-do. The slope of an area does not necessarily affect the probability of forest fires, but largely impacts the behavior of the forest fires. In other words, in areas where man-made fires account for most of the causes of forest fires, slope is one of the key factors of FFSM. This is also consistent with the results of previous studies in areas where most man-made forest fires were caused (Jaiswal et al. 2002). Approximately 70% of South Korea’s forest areas are located at a slope below 30° that is convenient not only for human activities but also for forest management. This indirectly leads to frequent human activities such as mountain climbing and camping in many mountainous areas near cities and streets. Moreover, as mentioned above, according to the statistics of Korea Forest Service in Gangwon Province, about 93% of forest fires in the study area were caused by fires created by people entering the mountains (incineration of garbage, cigarette fire, camping, etc.) (Korea Forest Service 2004, https://www.forest.go.kr). Correspondingly, the FFSM produced in this study reveals important information about the occurrence of forest fires in the study area, as it reflects the characteristics of human activities and interference in the study area (Inventory of forest fires data) and links the distance (proximity) factor. The probability of forest fires in low slope areas of Gangwon Province may also be very high. In this study, the FFSM established by considering the characteristics of regional topography and human activities was very beneficial in revealing the susceptible locations of forest fires in the research area and for planning the areas that need fire prevention.

As shown in Figure 10, the most susceptible (Very High) areas in the FFSM of Gangwon-do have slopes ranging from 0° (min) to 68.2° (max), with an average of 17.27° (mean). Therefore, one of the most important spatial characteristics in the occurrence of forest fires of forest fires in Gangwon-do is its low slope of

![Figure 10. Spatial analysis and relationship assessment between the very high susceptibility areas in the FFSM and the slope factor. Left: Spatial range of slope factors in the very high susceptibility area in FFSM in Gangwon-do; Right: Spatial analysis of slope by susceptibility level of FFSM.](image)

GEOMATICS, NATURAL HAZARDS AND RISK 445
approximately 17°. The region mostly covers adjoining urban and cropland areas, where human activity is high because of the low slope. The relationship between the FFSM and slope can be evaluated spatially (right portion of Figure 10), which would greatly help managers and planners in the prevention and management of the occurrence of forest fires and in the construction of bio-protection zones and ecological pathways in Gangwon-do.

5. Conclusion

Forest fires are one of the most destructive natural hazards in terms of property and ecosystem damage, causing substantial economic loss every year. Therefore, prediction and modeling of forest fires are essential, and forest fire susceptibility maps (FFSMs) can provide key information to planners and decision makers in the preparation of forest fire prevention and management plans.

In this study, a method for constructing an efficient and highly accurate FFSM for the Gangwon-do region of Korea was developed by combining the Google Earth Engine (GEE) platform and machine learning algorithms. After selecting 11 factors based on the existing research and the characteristics of the site, 70% of the forest fire inventory maps (270 samples) were trained using three machine learning algorithms (CART, RF, and BRT) to construct the FFSMs. The important factors were identified from 11 factors of the trained model. Accuracy was validated via the AUC value of the FFSMs using the remaining 30% of the dataset. The results of this study can be summarized as follows.

1. Among the three machine learning models, RF and BRT showed excellent prediction accuracy. The AUC value of BRT was 0.846, followed by the AUC value of RF of 0.835. BRT showed a more concentrated distribution in vulnerable areas. CART had the lowest AUC value (0.751), which shows the lowest prediction accuracy.
2. The FFSM calculated using BRT revealed that the factor with the greatest influence on forest fires is slope. Slope is one of the most important factors in forest fires because it greatly affects the fire spreading rate. Spatial analysis confirmed that a slope of 17° in Gangwon-do had a significant effect on the occurrence of forest fires.
3. The spatial analysis of FFSMs and the importance analysis of the factors confirmed the substantial impact of human activity and interference on forest fires in Gangwon-do.

The limitations of this study can be summarized as follows: First, it is difficult to acquire key data and establish important factors due to the specificity of the study area. The Gangwon-do region is close to the demilitarized zone (DMZ) between South Korea and North Korea, which makes it difficult to construct soil property factors throughout Gangwon-do. Second, it is difficult to exclude the influence of cloud mask and time scale in the construction of satellite image indices (e.g., NDVI) using satellite image data. Although temporal aggregation methods were used to construct
clean data that reduced the effect of clouds, it was difficult to construct the data for each forest fire occurrence period. Third, there is uncertainty in accuracy validation using only the ROC curve (AUC value). In machine learning, there is no clear answer to the question of which model is more accurate because the AUC value is affected by the selection of factors and the randomness of inventory maps. In this study, after accuracy verification using the ROC curve, visual interpretation and validation were performed in spatial terms by additionally considering the target research area and the characteristics of forest fires. In future studies, we will consider interrelating the fires with other disasters while maintaining the efficiency and high-accuracy of the forest fire analysis. We intend to separate the effects of human activities and interference on forest fires and evaluate the influencing factors required for the preparation of management plans.

Acknowledgements

Thanks to all anonymous reviewers and editors for taking the time to review this paper. Authors would also like to thankfully Google Earth Engine platforms (GEE) through which the data/products have been accessed and analyzed for the present study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Korea Environment Industry & Technology Institute(KEITI) through The Decision Support System Development Project for Environmental Impact Assessment, funded by Korea Ministry of Environment(MOE)(2021003360002).

Data availability statement

The data that support the findings of this study are available from the corresponding author, [Yong Piao], upon reasonable request.

References

Abedi Gheshlaghi H, Feizizadeh B, Blaschke T. 2020. GIS-based forest fire risk mapping using the analytical network process and fuzzy logic. J Environ Plan Manag. 63(3):481–499.
Amatulli G, Rodrigues MJ, Trombetti M, Lovreglio R. 2006. Assessing long-term fire risk at local scale by means of decision tree technique. J Geophys Res. 111(G4):G04S05.
Bae S-H. 2018. Spatial clustering analysis of fire in Gangwon-Do. J Korean Assoc Geogr Inf Stud. 21(3):93–103.
Breiman L, Friedman J, Stone CJ, Olshen RA. 1984. Classification and regression trees. Wadsworth Inc: CRC press.
Breiman L. 2001. Random forests. Mach Learn. 45(1):5–32.
Chuvieco E, Cocero D, Riaño D, Martin P, Martínez-Vega J, de la Riva J, Pérez F. 2004. Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating. Remote Sens Environ. 92(3):322–331.
Cochrane MA. 2001. Synergistic interactions between habitat fragmentation and fire in evergreen tropical forests. Conserv Biol. 15(6):1515–1521.

Cochrane MA. 2003. Fire science for rainforests. Nature. 421(6926):913–919.

Elith J, Leathwick JR, Hastie T. 2008. A working guide to boosted regression trees. J Anim Ecol. 77(4):802–813.

Eun-Jin P, Inae Y. 2018. Changes in policy and research regarding the Korean demilitarized zone and the implications. J Environ Policy Admin. 26(2):19–45. doi: http://dx.doi.org/10.15301/jepa.2018.26.2.19.

Fawcett T. 2006. An introduction to ROC analysis. Pattern Recognit Lett. 27(8):861–874.

Friedman JH. 2001. Greedy function approximation: a gradient boosting machine. Ann Statist. 29(5):1189–1232.

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens Environ. 202:18–27.

Hernandez-Leal P, Arbelo M, Gonzalez-Calvo A. 2006. Fire risk assessment using satellite data. Adv Space Res. 37(4):741–746.

Hu Y, Hu Y. 2019. Land cover changes and their driving mechanisms in Central Asia from 2001 to 2017 supported by Google Earth Engine. Remote Sens. 11(5):554.

Jaiswal RK, Mukherjee S, Raju KD, Saxena R. 2002. Forest fire risk zone mapping from satellite imagery and GIS. Int J Appl Earth Obs Geoinf. 4(1):1–10.

Keenan RJ, Reams GA, Achar F, de Freitas JV, Grainger A, Lindquist E. 2015. Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015. For Ecol Manag. 352:9–20.

Korea Forest Service. 2004. Forest location in Korea. Korea Forest Service. (In Korean).

Krasnow K, Schoennagel T, Veblen TT. 2009. Forest fuel mapping and evaluation of LANDFIRE fuel maps in Boulder County, Colorado, USA. For Ecol Manag. 257(7):1603–1612.

Kushla JD, Ripple WJ. 1997. The role of terrain in a fire mosaic of a temperate coniferous forest. For Ecol Manag. 95(2):97–107.

Lecina-Diaz J, Martínez-Vilalta J, Alvarez A, Banqué M, Birkmann J, Feldmeyer D, Vayreda J, Retana J. 2021. Characterizing forest vulnerability and risk to climate-change hazards. Front Ecol Environ. 19(2):126–133.

Lee S-Y, Lee H-P. 2006. Analysis of forest fire occurrence in Korea. Fire Sci Eng. 20(2):54–63.

Lentile LB, Holden ZA, Smith AM, Falkowski MJ, Hudak AT, Morgan P, Lewis SA, Gessler PE, Benson NC. 2006. Remote sensing techniques to assess active fire characteristics and post-fire effects. Int J Wildland Fire. 15(3):319–345.

Liu J, Duan Z. 2018. Quantitative assessment of landslide susceptibility comparing statistical index, index of entropy, and weights of evidence in the Shangnan area, China. Entropy. 20(11):868.

Liu X, Hu G, Chen Y, Li X, Xu X, Li S, Pei F, Wang S. 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. Remote Sens Environ. 209:227–239.

McKinley DC, Ryan MG, Birdsey RA, Giardina CP, Harmon ME, Heath LS, Houghton RA, Jackson RB, Morrison JF, Murray BC, et al. 2011. A synthesis of current knowledge on forests and carbon storage in the United States. Ecol Appl. 21(6):1902–1924.

Milanović S, Marković N, Pamučar D, Gigović L, Kostić P, Milanović SD. 2020. Forest fire probability mapping in Eastern Serbia: Logistic regression versus random forest method. Forests. 12(1):5.

Moayedi H, Mehrabi M, Bui DT, Pradhan B, Foong LK. 2020. Fuzzy-metaheuristic ensembles for spatial assessment of forest fire susceptibility. J Environ Manage. 260:109867.

Mollicone D, Eva HD, Achard F. 2006. Ecology: human role in Russian wild fires. Nature. 440(7083):436–437.

Nachappa TG, Ghorbanzadeh O, Gholamnia K, Blaschke T. 2020. Multi-hazard exposure mapping using machine learning for the state of Salzburg, Austria. Remote Sens. 12(17):2757.
Naghibi SA, Pourghasemi HR, Dixon B. 2016. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environ Monit Assess. 188(1):1–27.

Oliphant AJ, Thenkabail PS, Teluguntla P, Xiong J, Gumma MK, Congalton RG, Yadav K. 2019. Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30-m data using a random forest classifier on the Google Earth Engine Cloud. Int J Appl Earth Obs Geoinf. 81:110–124.

Park S-J, Lee D-K. 2020. Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. Environ Res Lett. 15(9):094052.

Piao Y, Jeong S, Park S, Lee D. 2021. Analysis of land use and land cover change using time-series data and random forest in North Korea. Remote Sens. 13(17):3501.

Pourghasemi HR, Beheshtirad M, Pradhan B. 2016. A comparative assessment of prediction capabilities of modified analytical hierarchy process (M-AHP) and Mamdani fuzzy logic models using Netcad-GIS for forest fire susceptibility mapping. Geomat Nat Haz Risk. 7(2):861–885.

Pourghasemi HR, Gayen A, Edalat M, Zarafshar M, Tiefenbacher JP. 2020. Is multi-hazard mapping effective in assessing natural hazards and integrated watershed management? Geosci Front. 11(4):1203–1217.

Pourghasemi HR, Gayen A, Lasaponara R, Tiefenbacher JP. 2020. Application of learning vector quantization and different machine learning techniques to assessing forest fire influence factors and spatial modelling. Environ Res. 184:109321.

Pourghasemi HR, Gayen A, Panahi M, Rezaie F, Blaschke T. 2019. Multi-hazard probability assessment and mapping in Iran. Sci Total Environ. 692:556–571.

Pourghasemi HR, Rahmati O. 2018. Prediction of the landslide susceptibility: which algorithm, which precision? Catena. 162:177–192.

Pourghasemi HR. 2016. GIS-based forest fire susceptibility mapping in Iran: a comparison between evidential belief function and binary logistic regression models. Scand J for Res. 31(1):80–98.

Pourtaghi ZS, Pourghasemi HR, Aretano R, Semeraro T. 2016. Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. Ecol Indic. 64:72–84.

Prosper-Laget V, Douguedroit A, Guinot J. 1995. Mapping the risk of forest fire occurrence using NOAA satellite information. EARSeL Adv Remote Sens. 4:30–38.

Satir O, Berberoglu S, Donmez C. 2016. Mapping regional forest fire probability using artificial neural network model in a Mediterranean forest ecosystem. Geomat Nat Haz Risk. 7(5):1645–1658.

Shabani S, Jaafari A, Bettinger P. 2021. Spatial modeling of forest stand susceptibility to logging operations. Environ Impact Assess Rev. 89:106601.

Shin JH, Lee DK. 2004. Strategies for restoration of forest ecosystems degraded by forest fire in Kangwon Ecoregion of Korea. For Ecol Manag. 201(1):43–56.

Sidhu N, Pebesma E, Câmara G. 2018. Using Google Earth Engine to detect land cover change: Singapore as a use case. Eur J Remote Sens. 51(1):486–500.

Siegent F, Ruecker G, Hinrichs A, Hoffmann A. 2001. Increased damage from fires in logged forests during droughts caused by El Niño. Nature. 414(6862):437–440.

Skole D, Tucker C. 1993. Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988. Science. 260(5116):1905–1910.

Tamimina H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, Brisco B. 2020. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS J Photogramm Remote Sens. 164:152–170.

Teodoro AC, Duarte L. 2013. Forest fire risk maps: a GIS open source application—a case study in Norwest of Portugal. Int J Geogr Inf Sci. 27(4):699–720.

Tutin CE, White LJ, Mackanga-Missandzou A. 1996. Lightning strike burns large forest tree in the Lope Reserve, Gabon. Glob Ecol Biogeogr Lett. 5(1):36–41.
Van Westen C, Rengers N, Soeters R. 2003. Use of geomorphological information in indirect landslide susceptibility assessment. Nat Haz. 30(3):399–419.

Verde J, Zêzere J. 2010. Assessment and validation of wildfire susceptibility and hazard in Portugal. Nat Hazards Earth Syst Sci. 10(3):485–497.

Woo H, Han S-W. 2020. Typological system of nature-based solutions and its similar concepts on water management. Ecol Resil Infrastruct. 7(1):15–25.