forests

Article

Can Ensemble Techniques and Large-Scale Fire Datasets Improve Predictions of Forest Fire Probability Due to Climate Change?—A Case Study from the Republic of Korea

Hyeon Kwon Ahn [1,2], Huicheul Jung [3] and Chul-Hee Lim [1,4,*]

[1] Department of Forestry, Environment, and Systems, Kookmin University, 77 Jeongneungro, Seongbukgu, Seoul 02707, Republic of Korea; postmist@kookmin.ac.kr
[2] Department of Forest Resources, Kookmin University, 77 Jeongneungro, Seongbukgu, Seoul 02707, Republic of Korea
[3] Korea Adaptation Center for Climate Change, Korea Environment Institute, Sejong 30147, Republic of Korea; hchjung@kei.re.kr
[4] College of General Education, Kookmin University, 77 Jeongneungro, Seongbukgu, Seoul 02707, Republic of Korea
* Correspondence: clim@kookmin.ac.kr

Abstract: The frequency of forest fires worldwide has increased recently due to climate change, leading to severe and widespread damage. In this study, we investigate potential changes in the fire susceptibility of areas in South Korea arising from climate change. We constructed a dataset of large-scale forest fires from the past decade and employed it in machine learning models that integrate climatic, socioeconomic, and environmental variables to assess the risk of forest fires. According to the results of these models, the eastern region is identified as highly vulnerable to forest fires during the baseline period, while the western region is classified as relatively safe. However, in the future, certain areas along the western coast are predicted to become more susceptible to forest fires. Consequently, as climate change continues, the risk of domestic forest fires is expected to increase, leading to the need for proactive prevention measures and careful management. This study contributes to the understanding of forest fire occurrences under diverse climate scenarios.

Keywords: forest fire; climate change; machine learning; ensemble; South Korea

Citation: Ahn, H.K.; Jung, H.; Lim, C.-H. Can Ensemble Techniques and Large-Scale Fire Datasets Improve Predictions of Forest Fire Probability Due to Climate Change?—A Case Study from the Republic of Korea. *Forests* **2024**, *15*, 503. https://doi.org/10.3390/f15030503

Academic Editor: Margarita Arianoutsou

Received: 9 January 2024
Revised: 17 February 2024
Accepted: 19 February 2024
Published: 8 March 2024

1. Introduction

Climate change has led to a dramatic increase in the frequency and intensity of hydroclimatic disasters (e.g., drought, flooding, and wildfire) over the past decade. For example, in 2022, Pakistan experienced extensive flooding over approximately one-third of its land area, while the 2019 Australian wildfires affected over 19 million ha of land, leading to ecological damage and substantial economic losses [1–3]. South Korea has also experienced numerous hydroclimatic disasters in recent years, including the 2019 Goseong Gangneung Forest Fire, the 2022 Samcheok–Uljin Forest Fire, concentrated heavy rainfall in the Gangwon region and the Seoul metropolitan area, and drought in the Yeongnam and Chungcheong regions, all of which have had a significant impact on both the human and natural environment [4].

With the greater prevalence of hot and dry weather conditions due to climate change, there is a worldwide increase in the frequency of wildfires, accompanied by a growing scale of damage [5–8]. Forest fires can lead to deforestation, habitat destruction, a reduction in biodiversity, and the loss of soil nutrients within forest ecosystems [9]. They can also have severe economic impacts, including decreased timber and livestock yields, the loss of environmental functions in forests, damage to national parks, industrial disruptions, and transportation disturbances [10]. The increasing frequency of forest fires is thus anticipated to severely disturb forest ecosystems in the long term [11,12].

To address forest fires, South Korea has implemented a range of policies. For example, the 6th Basic Forest Plan aims for a 21% increase in forest growing stock by 2050 compared to 2015. However, this plan acknowledges the persistent risk of large-scale forest fires and resulting damage as residential areas expand into adjacent forested regions. The plan thus aims to proactively prevent and respond to forest fires by employing advanced information technology to minimize both the occurrence and impact of forest fires [13]. Nevertheless, the 6th Basic Forest Plan acknowledges the risks associated with expanding residential areas and the potential for large-scale wildfires and their damages in the current firefighting policies. However, it lacks the necessary provisions for predicting future forest fire occurrences. Forest fire prediction is a vital component of forest fire management in terms of assessing the risk of forest fires, strengthening forest fire monitoring and suppression efforts, and facilitating strategic planning and resource allocation for firefighting [14,15].

Among the various methods for predicting forest fires, machine learning is widely employed. Machine learning, a branch of artificial intelligence (AI) and computer science, focuses on utilizing data and algorithms to replicate the iterative learning process of humans, thereby improving accuracy over time [16,17]. Major machine learning algorithms include Maximum Entropy (MaxEnt), Gradient Boosting Machine, and Artificial Neural Networks (ANNs). Maximum Entropy is an optimization technique that determines the most probable probability distribution given specific constraints, proving effective in predicting wildfire probabilities. Gradient Boosting Machines utilize ensemble learning, enhancing prediction performance by amalgamating multiple decision trees. Artificial Neural Networks, resembling networks of neural cells, serve as learning algorithms capable of modeling complex nonlinear relationships. Additionally, a number of past studies have built models based on machine learning techniques for the prediction of forest fires and have used their output to assess forest fire risks [18–20]. Furthermore, with the escalating risk of large-scale forest fires, it is imperative to formulate a forest fire prediction model specifically tailored to anticipate and mitigate the impacts of such extensive occurrences, aiming to minimize the resultant damages.

This study forecasts the likelihood of forest fires considering Shared Socioeconomic Pathways (SSP) scenarios. By incorporating climate, topography, environmental, socioeconomic, and historical forest fire locations into machine learning and statistics-based models, we aim to predict the probability of medium–large-scale forest fires by considering SSP scenarios using an ensemble technique. This approach seeks to offer scientific insights into the occurrence and response mechanisms of forest fires under emerging climate change scenarios. In contrast to prior studies that have focused on predicting current forest fire risks based on existing climate conditions, our methodology enables the assessment of future forest fire risks. This research contributes to the evaluation of forest fire risks in South Korea and the formulation of anticipatory measures for the future forest ecosystem.

2. Data and Methods

2.1. Study Area

This study was conducted in South Korea, which is located at 33–39° N, 124–131° E, with a total area of approximately 100,443.6 km² [21]. Forest covers approximately 62.7% of the total land area, of which approximately 36.9% is classified as vulnerable deciduous forest, which is a relatively high proportion [22]. The forest density increased due to the implementation of the 1st and 2nd Forest Basic Plan projects from 1973 to 1987, creating an environment that is more susceptible to large-scale forest fires, especially with the higher proportion of mature forests (21–50 years old) [23]. The annual precipitation in South Korea is approximately 1306.33 mm.

Most of this rainfall falls during summer, with approximately 58.4% of the annual precipitation occurring during this season in 2022 [24]. As climate change worsens and the aging of forests continues, an increase in forest fires in South Korea is predicted [7].

2.2. Data

In the present study, the analysis period was divided into baseline years (2010–2019), the near future (2040–2049), and the distant future (2070–2079) to track the potential impact of climate change in accordance with the Basic Forest Plan. To establish the forest fire prediction model, we collected information on medium- to large-scale forest fires recorded from 2010 to 2019 provided by the Korea Forest Service to predict forest fires caused by climate change (Figure 1). According to the regulations outlined in the "Regulations on the Mission and Role of Forest Firefighting Agencies" by the Korea Forest Service, a large-scale forest fire is one that spreads over an area of 100 ha or more, while a medium-scale forest fire refers to cases that do not meet the criteria for a large-scale forest fire. In this study, a medium-scale forest fire was thus defined as a forest fire with a damage area of 4 ha or more, and the prediction of the potential occurrence of significant forest fires of medium–large scale over a period of 10 years was referred to as the forest fire probability (FFP).

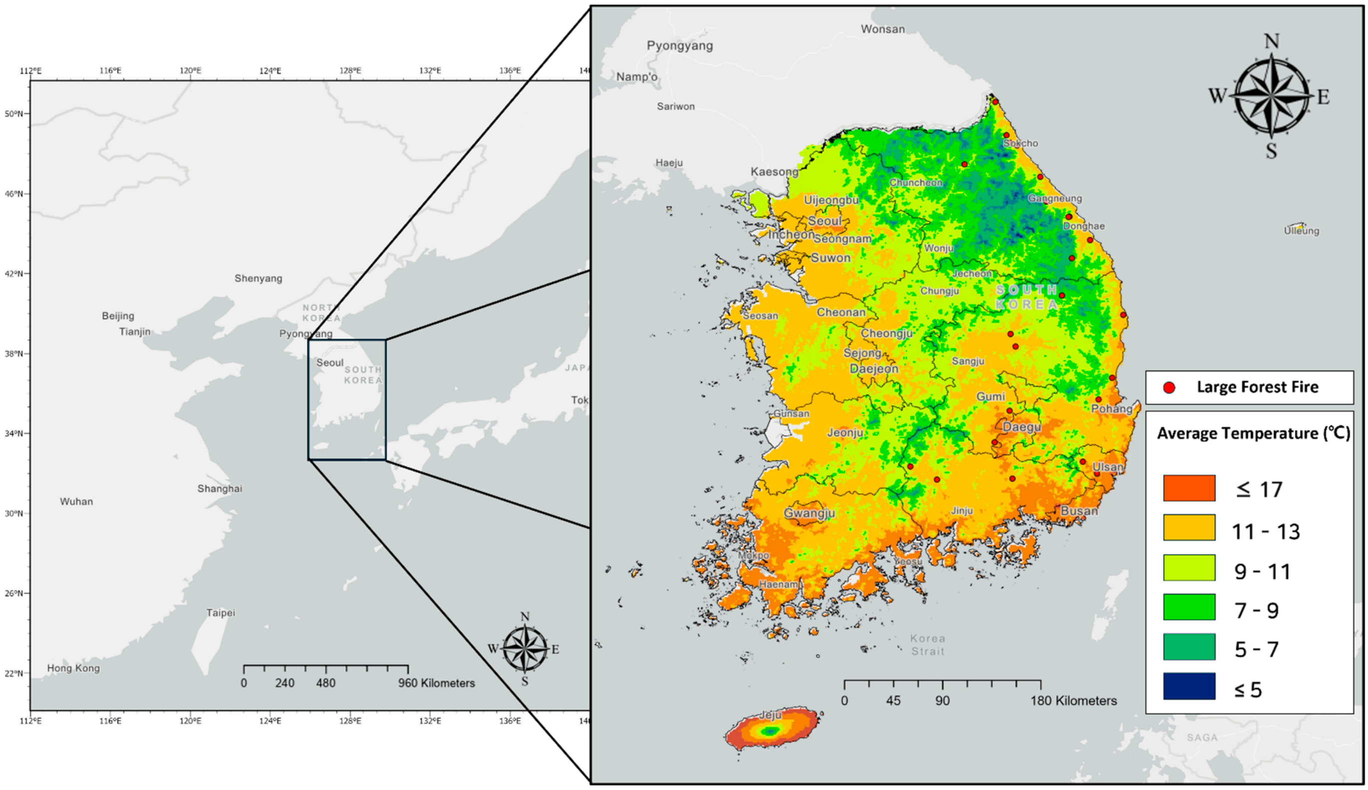

Figure 1. Study area with average temperature (source: Korea Meteorological Administration) and large forest fire occurrence points (source: Korea Forest Service).

Climatic, topographical, socioeconomic, and environmental variables were employed in the individual and ensemble models to predict the FFP (Table 1). The climate data utilized for this analysis had a spatial resolution of 1 km² and included the highest monthly maximum temperature, lowest monthly minimum temperature, average relative humidity in spring (March–May), precipitation during the driest quarter, and maximum wind speed [25]. Climate information for the baseline period was derived from MK-PRISM v2.1 data produced by the Korea Meteorological Administration (KMA), while climate information for the near and distant future was based on the SSP climate change scenarios provided by the KMA [26]. PRISM, developed by Oregon State University in the United States, is a model that statistically refines the observation data from automatic weather stations and automated synoptic observing systems to produce high-resolution grid-based climate data, incorporating factors such as the slope, distance, elevation, and marine influences [27–29]. MK-PRISM v2.1 is a modified Korean version of PRISM that provides a high-resolution grid-based climate dataset containing daily, monthly, and annual data from 2000 to 2019 with a spatial resolution of 1 km.

Table 1. Input variables used to estimate forest fire probability models.

Input Variable	Variable Name	Unit	Source
Climate	Max. temperature of the warmest month	°C	Korea Meteorological Administration
	Min. temperature of the coldest month	°C	Korea Meteorological Administration
	Precipitation of the driest month	mm	Korea Meteorological Administration
	Max. windspeed	m/s	Korea Meteorological Administration
	Relative humidity during spring	%	Korea Meteorological Administration
Topographic	Elevation	m	NASA
	Topographic wetness index	-	-
	Aspect	-	-
Socio-economic	Distance from urban area	m	Land cover map from the Korea Ministry of Environment
	Population density	People per km^2	Population census from the Korean Statistical Information Service
Environment	Forest type	-	Land cover map from the Korea Ministry of Environment

The Representative Concentration Pathway scenarios utilized in the IPCC 5th Assessment Report were designed based on the radiative forcing exerted by human activity on the atmosphere and provide greenhouse gas concentration trajectories. In contrast, the SSP scenarios introduced in the 6th Assessment Report incorporate potential social and economic changes arising from mitigation and adaptation efforts in response to climate change. The SSP scenarios are thus more realistic in terms of future projections. The data for the SSP scenarios originate from the UK Met Office's prediction model UKESM1. Utilizing data from UKESM1, an ensemble average is calculated by dynamically downscaling into regional climate models, generating climate change scenarios with a spatial resolution of 25 km. Subsequently, each model is refined using PRISM, and an ensemble is constructed for the five models to generate a climate change scenario at a spatial resolution of 1 km [30].

In this study, three SSP scenarios were employed. SSP1-2.6 assumes a sustainable development path in which the rational and rapid progression of sustainable development minimizes fossil fuel use due to the development of renewable energy. SSP2-4.5 represents a scenario with moderate levels of mitigation and adaptation efforts for climate change, while SSP5-8.5 assumes high fossil fuel use and greater urban-centric development. The SSP1-2.6 scenario is particularly relevant in this context because the IPCC has recommended achieving global carbon neutrality by 2050 to limit the global average temperature increase to 1.5 degrees by 2100. Therefore, SSP1-2.6 is considered the scenario most similar to carbon neutrality.

For the elevation, terrain wetness index (TWI), and aspect, which were employed as the topographical variables, we utilized a digital elevation model with a spatial resolution of 1 km provided by the Shuttle Radar Topography Mission (SRTM). The TWI was used to reflect the soil moisture levels, which influence the dryness of the leaf litter and understory vegetation, which are susceptible to ignition during a forest fire. Aspect represents the direction of the slope within a grid; in the Northern Hemisphere, south-facing slopes receive more sunlight, leading to drier conditions and a higher forest fire risk compared to north-facing slopes. Empirical evidence from a study conducted during the 2007–2009 spring seasons that tracked a total of 101 forest fires in South Korea revealed that the forest fire frequency on south-facing slopes was approximately 2.7 times higher than on north-facing slopes [31].

The socioeconomic variables employed in the analysis were the distance from urban area taken from land cover data and the population density per unit area. Social factors that have been identified as having a significant correlation with forest fire occurrences also have a positive correlation with population density in the modeling process [32], while the distance from urban area, calculated using the Euclidean distance method, represents the accessibility to forests. The Euclidean distance is utilized to measure the shortest distance between two points. This measurement method can also be applied in multidimensional space and is employed to calculate the similarity or distance between data points [33]. These variables were included because approximately 178.3 cases of unintentional ignition and 30.4 cases of cigarette-caused ignition occur on average each year, highlighting the role of human activity in forest fires. Environmental variables were employed to reflect the vulnerability to forest fires based on vegetation type, differentiating between coniferous, deciduous, and mixed forests.

To enhance the accuracy of the modeling process and mitigate multicollinearity between variables, Pearson correlation coefficient analysis was conducted [34,35].

Factors with correlation coefficients exceeding 0.8 were excluded from the models. However, despite the negative correlation of –0.81 between the maximum temperature of the warmest month and elevation, elevation was retained in the modeling process because elevation incorporates important topographical elements and anthropogenic factors such as climate and accessibility.

2.3. Machine Learning and Ensemble Method

In this study, machine learning algorithms were employed to predict the occurrence of forest fires based on climatic variables, topographic factors, socioeconomic, and environmental variables. The machine learning algorithms used for the analysis included Random Forest (RF), Maximum Entropy (MaxEnt), Generalized Boosting Model (GBM), Artificial Neural Network (ANN), Classification Tree Analysis (CTA), and Flexible Discriminant Analysis (FDA), while the statistical-based Multivariate Adaptive Regression Splines (MARS) and Generalized Linear Model (GLM) were also used. Except for the MaxEnt model, the other models required both occurrence and non-occurrence data for forest fire locations, with a 5:5 ratio maintained via random sampling.

The accuracy of both the individual models and the ensemble model was evaluated using the area under the curve (AUC) derived from the receiver operating characteristic (ROC) curve. The verification of the model's explanatory power using the AUC was utilized to assess the diverse models. It is deemed that the closer the AUC value is to 1, the more accurately the model predicts outcomes. If the AUC value exceeds 0.7, the model is recognized to possess substantial explanatory power [36,37]. Model validation involved splitting the dataset into training and testing data at an 8:2 ratio.

To enhance the accuracy of the ensemble model, the eight models were repeated 10 times, and an ensemble was run based on the mean ensemble approach, with only the results with an AUC of greater than 0.7 considered. The ensemble method runs multiple models, combines their predictive values, and utilizes them for final decision making (Figure 2). This approach has been widely used in various studies to compensate for the errors and uncertainties of individual models [38–42].

The output of both the single models and the ensemble was an FFP ranging from 0 to 1. The FFP results were categorized into five grades using equal intervals. Areas with an FFP of 0.2 or lower were classified as safe from forest fires, those with an FFP of 0.2–0.4 were considered relatively safe, those with an FFP of 0.4–0.6 were considered to have a relatively low fire risk, those with an FFP of 0.6–0.8 were considered to have a relatively high fire risk, and those exceeding 0.8 were classified as high-risk areas.

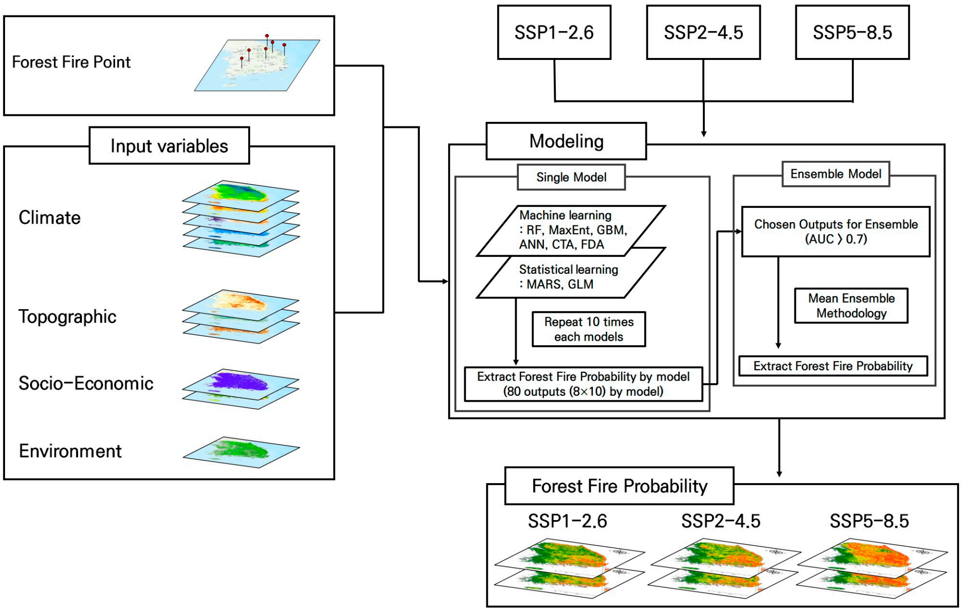

Figure 2. Conceptual and methodological diagram of our study.

3. Results and Discussion

3.1. Model Accuracy and Variable Importance

The average accuracy from 10 repeated runs for each model is presented in Table 2. The accuracy of the ANN and CTA models was relatively low compared with the MaxEnt, RF, and GBM models. Models with AUC values exceeding 0.7 were combined in the ensemble, which consequently exhibited a higher AUC than the individual models, confirming its effectiveness in mitigating the uncertainties and errors associated with single-model approaches.

Table 2. Area under the curve (AUC) values for each model.

Model		AUC
Machine learning	Random Forest	0.732
	MaxEnt	0.746
	Generalized Boosting Model	0.725
	Artificial Neural Networks	0.590
	Classification Tree Analysis	0.579
	Flexible Discriminant Analysis	0.677
Statistical learning	Multivariate Adaptive Regression Splines	0.689
	Generalized Linear Model	0.660
Ensemble	Mean	0.976

The importance of factors influencing the forest fire prediction models is presented in Table 3. The method of evaluating importance proceeds as follows: Initially, model prediction is performed using a dataset in which a single variable is randomly mixed from the given data. Then, the Pearson correlation coefficient between the reference prediction and the prediction derived from the dataset is calculated. After that, the larger the value obtained by subtracting the correlation coefficient from 1, the greater the influence of the variable on the model, and a value close to 0 indicates that the influence of the variable on the model is small [43]. The importance is presented as a relative value, averaging the significance of each factor across 10 repetitions of the eight models. Elevation had the highest importance (0.352), followed by relative humidity during spring, minimum temperature of the coldest month, and aspect.

Factors aside from altitude are highly correlated with fuel dryness, underscoring that the dryness of the fuel is a substantial contributing factor to forest fire incidents. This underscores the significance of precipitation in formulating forest fire prediction and response strategies.

Table 3. Mean importance of the variables for the individual models.

Variable Type	Variable	Mean Importance
Climate	Max. temperature of the warmest month	0.066
	Min. temperature of the coldest month	0.175
	Precipitation of the driest month	0.094
	Max. windspeed	0.021
	Relative humidity during spring	0.268
Topographic	Elevation	0.352
	Topographic Wetness Index	0.036
	Aspect	0.147
Socio-Economic	Distance from used area	0.026
	Population density	0.064
Environment	Forest type	0.030

3.2. Forest Fire Probability Attributable to SSP Scenario

3.2.1. Baseline Period

The FFP predictions made by the ensemble model for the 2010–2019 period are presented in Figure 3. During this baseline period, regions along the eastern coast and some inland areas exhibited relatively high FFP, resembling the characteristics of major domestic forest fires. Gangwon-do, Gyeongsang-do, and the Seoul metropolitan area were relatively vulnerable to forest fires, while regions along the western were relatively safe from forest fires. The high FFP in Gangwon-do was attributed to strong winds, steep terrain, and the presence of vulnerable pine forests, meaning that this region had a significant risk of major forest fires breaking out [44]. Moreover, in the western region, there is a higher prevalence of broad-leaved forests in comparison to coniferous forests, and these forests exhibit a relatively fragmented distribution. Furthermore, the non-forest area is more extensive than that in the eastern coast region, contributing to the comparatively lower vulnerability of the west coast area compared to the eastern region.

To evaluate the accuracy of the forest fire susceptibility model, a comparative analysis was conducted by examining the occurrence of large-scale forest fires over the past 20 years. In the last two decades, a total of 24 major forest fires have been recorded, 21 of which occurred in areas with an FFP of over 0.8. The ensemble model thus demonstrated high reproducibility for large-scale forest fires, indicating that it effectively learns climatic variables based on training using accumulated occurrence data, particularly in relation to major forest fires.

3.2.2. SSP1-2.6

The FFP predictions made by the ensemble model for SSP1-2.6 are presented in Figure 4. Similar to the baseline period, the eastern region near the Taebaek Mountains was identified as the area most prone to forest fires. Due to efforts to mitigate climate change, the forest fire risk in Gyeongsangbuk-do decreased in both the near and distant future compared to the baseline period. However, in the near future, the forest fire risk in the western region increased compared to the baseline period before subsequently weakening over time due to the gradual reduction in greenhouse gas levels. The gradual rise in FFPs in the western region implies a growing likelihood of extensive wildfires in broad-leaved forests. This indicates that even presently considered relatively secure broad-leaved forests may become susceptible to forest fires with the progression of climate change.

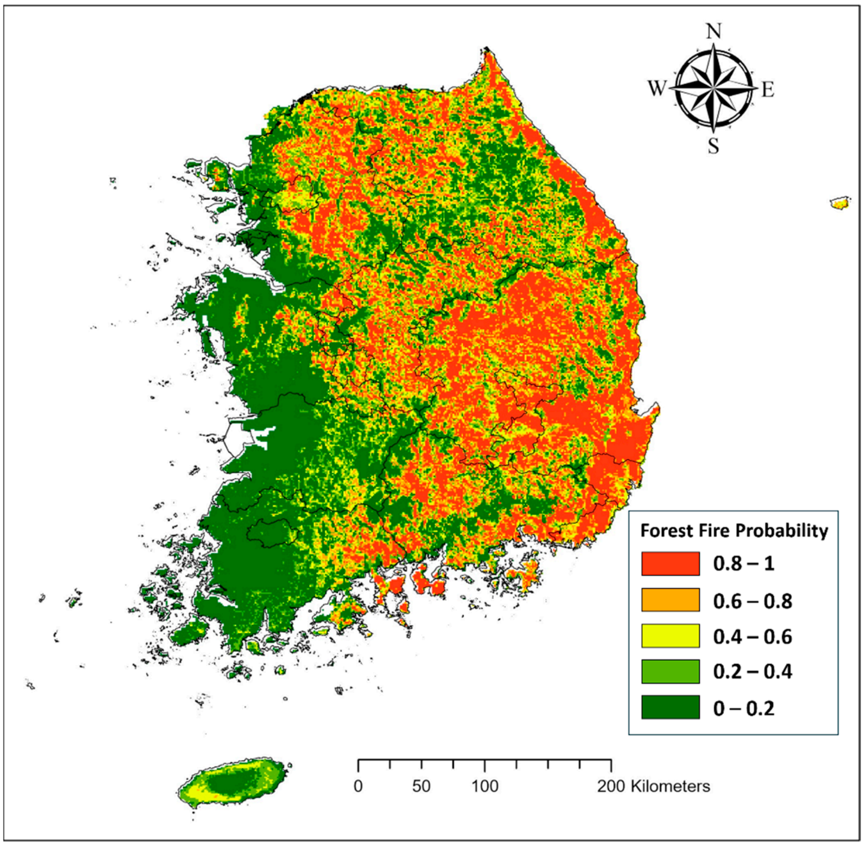

Figure 3. Baseline period forest fire probability (FFP) according to ensemble model.

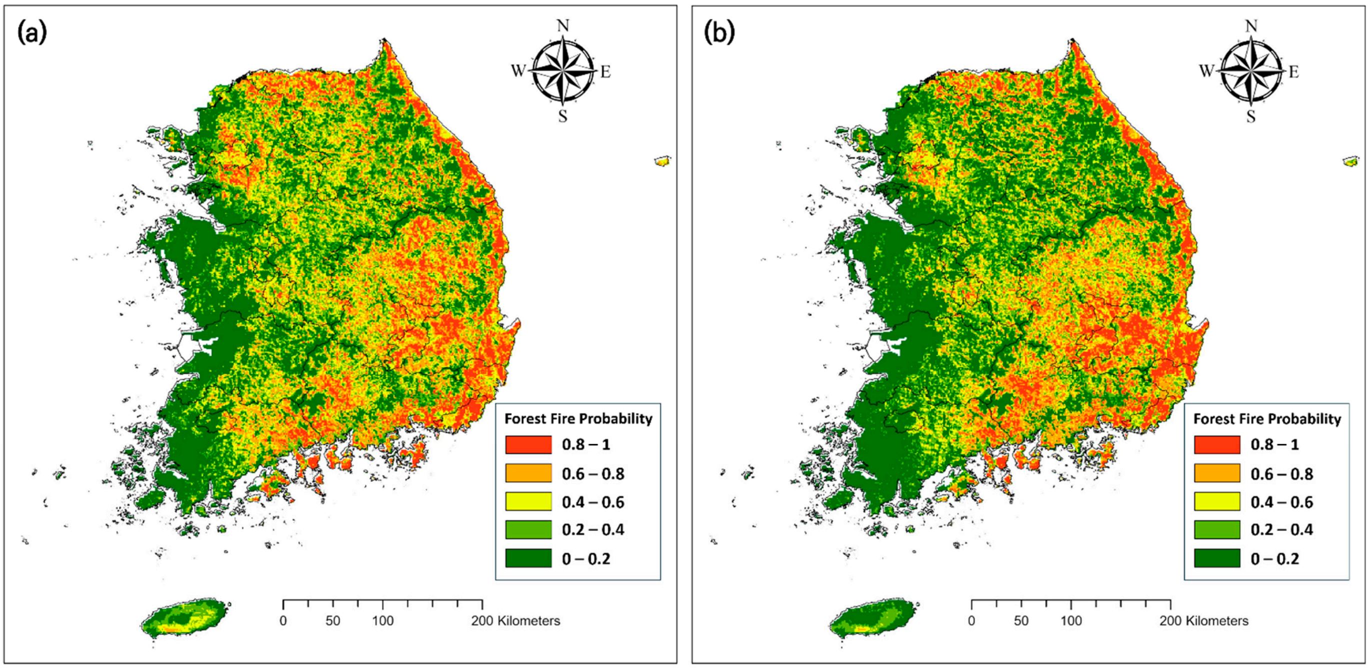

Figure 4. SSP1-2.6 FFP according to ensemble model: (**a**) 2040s FFP; (**b**) 2070s FFP.

3.2.3. SSP2-4.5

The FFP predictions made by the ensemble model for SSP2-4.5 are displayed in Figure 5. As with the baseline and the SSP1-2.6 scenario, the eastern region near the Taebaek Mountains was identified as particularly vulnerable to forest fires. As time progresses from the near future to the distant future, the risk of forest fires in Gyeongsangbuk-do significantly escalates. The topography of South Korea generally exhibits an elevation increase from west to east, and this is anticipated to contribute to the heightened FFP in high-altitude areas as climate change advances. While the size of the area with a high FFP compared to the baseline year was generally lower, there were more forest fire-prone areas in the western region. However, unlike the SSP1-2.6 scenario, the stronger forest fire risk in the western region was projected to continue without a significant reduction in the far future.

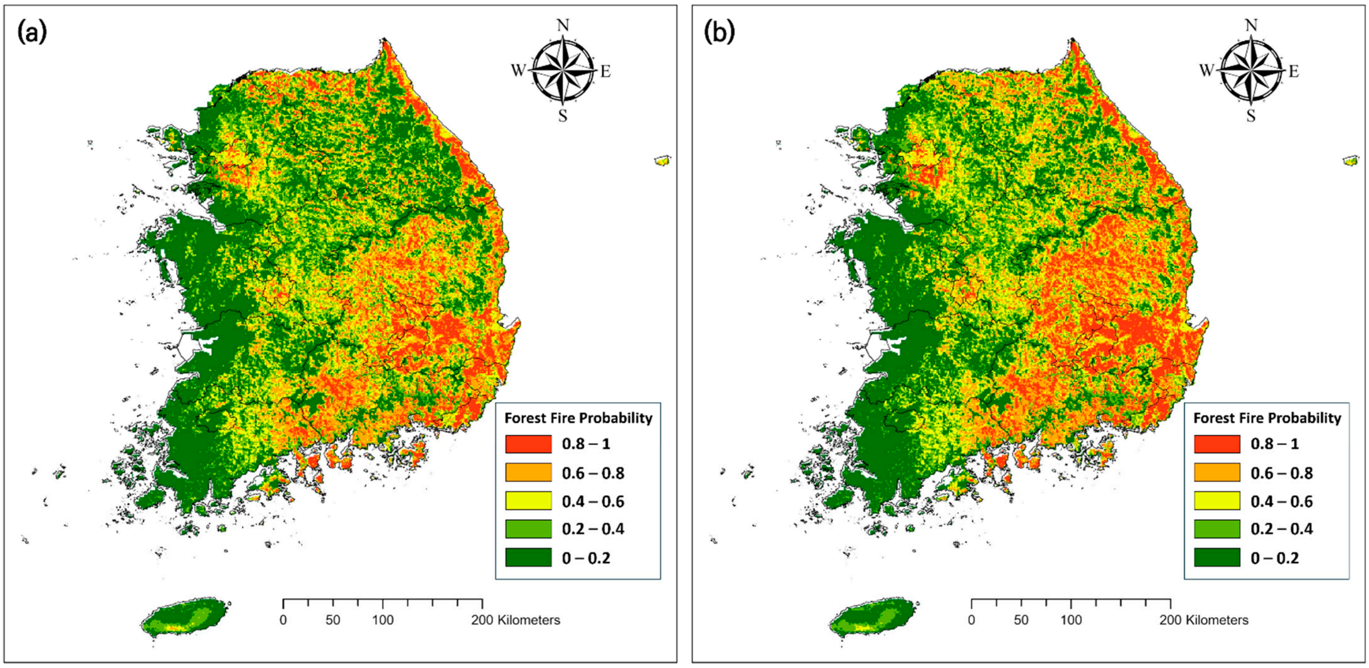

Figure 5. SSP2-4.5 FFP according to ensemble model: (**a**) 2040s FFP; (**b**) 2070s FFP.

3.2.4. SSP5-8.5

The FFP predictions made by the ensemble model for SSP5-8.5 are presented in Figure 6. The SSP5-8.5 scenario led to an increase in forest fire-prone areas compared to the existing scenarios, with a relatively high FFP observed in most central and eastern regions, excluding some high-elevation areas in Gangwon-do. The greater the elevation of mountain ranges in South Korea, the more formidable the terrain becomes, impeding the movement of manpower and presenting challenges in extinguishing forest fires. This is primarily due to the limited availability of water for firefighting purposes. According to our findings derived from considering the SSP5-8.5 scenario, which exacerbates vulnerability to wildfires in numerous regions, it is anticipated that small-scale forest fires may escalate into larger-scale forest fires.

In contrast to the SSP1-2.6 and SSP2-4.5 scenarios, the SSP5-8.5 scenario led to a continued rise in forest fire-prone areas in the near future, surpassing those of the baseline period. Forest fire-prone areas were also expected to expand in various regions, including Gyeongbuk, the Seoul metropolitan area, and Daejeon, thus differing from the trends observed for the other scenarios.

3.2.5. Changes in Forest Fire Risk

To assess the forest fire risk for the different SSP scenarios, regions with an FFP of 0.6 or higher were identified as areas susceptible to forest fires. The total area classified as susceptible to forest fires for each SSP scenario is presented in Table 4. In comparison to the baseline period, the area susceptible to forest fire decreased for the SSP1-2.6 and SSP2-4.5 scenarios, while it was higher for the SSP5-8.5 scenario. However, although the

SSP2-4.5 scenario exhibited a reduction in forest fire risk in the near future, a subsequent rise was projected for the distant future.

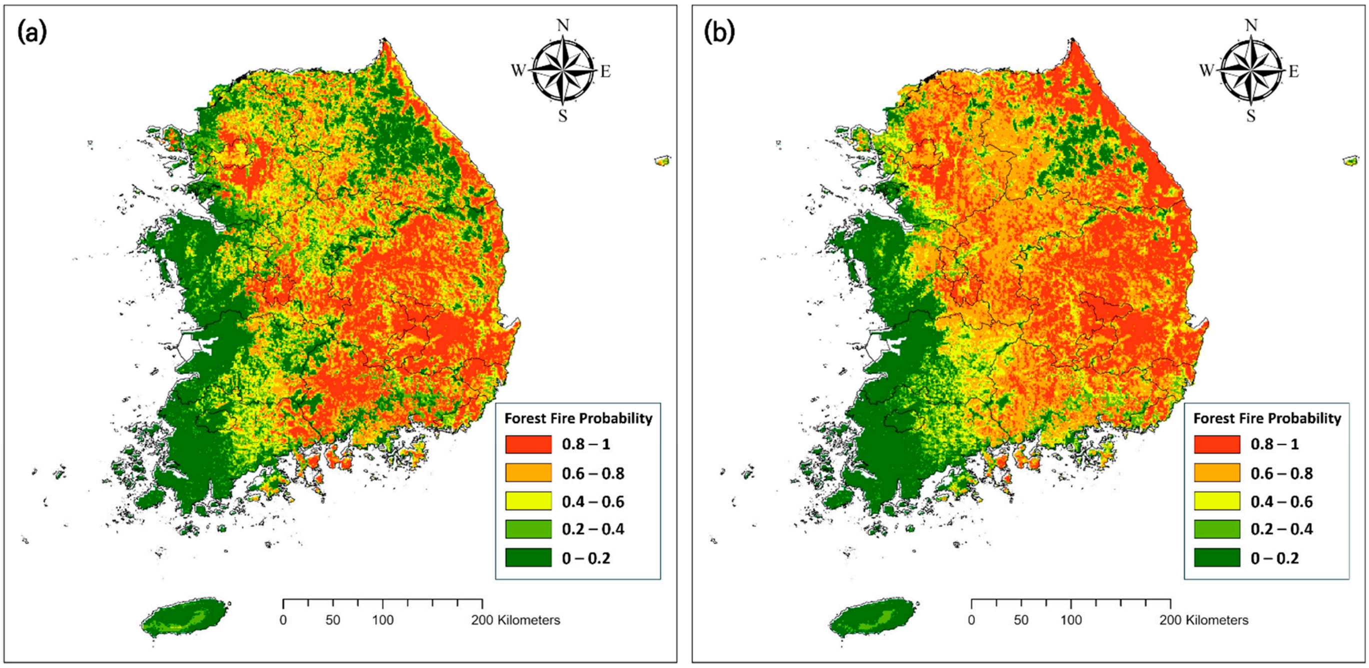

Figure 6. SSP5-8.5 FFP according to ensemble model: (**a**) 2040s FFP; (**b**) 2070s FFP.

Table 4. Forest fire risk area by SSP scenario (unit: km^2).

Model	Current	2040s			2070s		
		SSP1-2.6	SSP2-4.5	SSP5-8.5	SSP1-2.6	SSP2-4.5	SSP5-8.5
Forest fire risk area	39,121	28,579	27,111	42,260	25,387	31,274	56,861

Considering the variations in forest fire risk areas associated with these scenarios, there is a compelling need for the continuous monitoring of forest fires and the enhancement of early extinguish systems to mitigate the escalation of localized incidents into large-scale forest fires. It is imperative to accurately anticipate high-risk forest fire areas in the future and devise suitable responses and preventive measures for these regions. Furthermore, it is crucial to establish a flexible and sustainable forest management policy based on the prediction of future forest fire risks under each scenario.

3.3. Implications and Limitations

In this study, an ensemble of various machine learning models trained using climatic, topographic, and demographic data pertaining to the locations of major forest fires was employed to mitigate the errors and uncertainties associated with individual fire risk prediction models. The AUC of the ensemble model was particularly high, standing at 0.976, surpassing the AUC of previous forest fire prediction models [45–48]. In addition, by training the ensemble using extreme climatic conditions associated with the spread of large forest fires following ignition, the FFP predictions for significant forest fires were shown to be useful for analyzing climate change adaptation measures.

However, there are limitations to this study. Forest fires are significantly influenced by socioeconomic factors. However, this study did not consider future changes in important socioeconomic variables in the modeling process. Furthermore, the analysis did not account for changes in environmental variables such as shifts in forest types, thus limiting the generalizability of the study. Additionally, topographic factor changes at a spatial resolution of 1 km are unpredictable and therefore could not be accounted for.

Particularly, machine learning models are significantly impacted by the training data they receive. In this study, the spatial coordinates of forest fire occurrences were derived from a limited timeframe, resulting in a concentration of coordinates primarily in the eastern region. The influence of these spatial coordinates is a contributing factor to the elevated FFP in the eastern region. To address these limitations in future studies, enhancing the dataset by extending the temporal coverage of forest fire occurrences could provide a more comprehensive understanding. Furthermore, factors crucial to the model's performance are those associated with fuel dryness. Improved results could be achieved by effectively incorporating measures that reflect the dryness of the fuel.

4. Conclusions

In recent years, hydrological disasters have become more frequent due to climate change, and numerous studies have focused on predicting and managing forest fires. Forest fire prediction plays a crucial role in forest fire management, with future occurrences being predicted based on historical or present climate conditions. In the present study, the FFP in Korea was assessed using an ensemble machine learning model incorporating new climate scenarios, topographic data, and socioeconomic data. Changes in the FFP for the near and distant future were then analyzed for different SSP scenarios.

In the baseline period, 39,121 km^2 of South Korea was classified as vulnerable to forest fires, representing approximately 38.9% of the country. However, in the 2070s and under the SSP5-8.5 scenario, this was projected to increase by approximately 34.7%. The eastern region was identified as highly vulnerable to forest fires in the baseline period, while the western region was classified as relatively safe. However, in the 2070s and under the SSP5-8.5 scenario, some of these safe areas in the western regions were predicted to become vulnerable to forest fires. In conclusion, as climate change continues, the risk of domestic forest fires is expected to rise, highlighting the need to develop effective preventative and management measures. This forest fire prediction model, developed utilizing forest fire occurrence data and diverse datasets, is positioned to offer foundational information to facilitate forest fire management and adaptation to climate change. Its deployment is expected to make a substantial contribution to the establishment of a safer environment and reduce damages resulting from forest fires.

Author Contributions: Conceptualization, methodology, investigation, and writing, H.K.A. and C.-H.L.; project administration, writing—review and editing, C.-H.L. and H.J. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is based on the findings of the research project "Forest vulnerability to climate change" (2023-029(R)), conducted by the Korea Environment Institute (KEI) and funded by Korea Forestry Promotion Institute (2022462C10-2324-0201). It was also funded by the National Research Foundation of Korea grant provided by the Ministry of Science and ICT (No. 2022R1C1C1008489).

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from Korea Forest Service and are available from the authors with the permission of Korea Forest Service.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Nanditha, J.S.; Kushwaha, A.P.; Singh, R.; Malik, I.; Solanki, H.; Chuphal, D.S.; Dangar, S.; Mahto, S.S.; Vegad, U.; Mishra, V. The Pakistan flood of august 2022: Causes and implications. *Earth's Future* **2023**, *11*, e2022EF003230. [CrossRef]
2. Nolan, R.H.; Boer, M.M.; Collins, L.; de Dios, V.R.; Clarke, H.; Jenkins, M.; Kenny, B.; Bradstock, R.A. Causes and consequences of eastern Australia's 2019–20 season of mega-fires. *Glob. Chang. Biol.* **2020**, *26*, 1039–1041. [CrossRef]
3. Ehsani, M.R.; Arevalo, J.; Risanto, C.B.; Javadian, M.; Devine, C.J.; Arabzadeh, A.; Venegas-Quiñones, H.L.; Dell'Oro, A.P.; Behrangi, A. 2019–2020 Australia Fire and Its Relationship to Hydroclimatological and Vegetation Variabilities. *Water* **2020**, *12*, 3067. [CrossRef]
4. Yeom, C.-H.; Lee, S.Y.; Park, H.S.; Kwon, C.G. A Study on Facilities Damage Characteristics Cused by Forest Fire in Goseong-Gun. *J. Korean Soc. Disaster Inf.* **2019**, *15*, 469–478.

5. Pechony, O.; Shindell, D.T. Driving forces of global wildfires over the past millennium and the forthcoming century. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19167–19170. [CrossRef]
6. Sutanto, S.J.; Vitolo, C.; Di Napoli, C.; D'Andrea, M.; Van Lanen, H.A. Heatwaves, droughts, and fires: Exploring compound and cascading dry hazards at the pan-European scale. *Environ. Int.* **2020**, *134*, 105276. [CrossRef]
7. Sung, M.K.; Lim, G.H.; Choi, E.H.; Lee, Y.Y.; Won, M.S.; Koo, K.S. Climate change over Korea and its relation to the forest fire occurrence. *Atmosphere* **2010**, *20*, 27–35.
8. Chung, Y.S.; Yoon, M.B.; Kim, H.S. On Climate Variations and Changes Observed in South Korea. *Clim. Change* **2004**, *66*, 151–161. [CrossRef]
9. Durán, R.; Varela, E.; Garza, D.; Prada, A.; Vázquez, M.; Soliño, M. Valuation of terrestrial and marine biodiversity losses caused by forest wildfires. *J. Behav. Exp. Econ.* **2017**, *71*, 88–95. [CrossRef]
10. Kalogiannidis, S.; Chatzitheodoridis, F.; Kalfas, D.; Patitsa, C.; Papagrigoriou, A. Socio-Psychological, Economic and Environmental Effects of Forest Fires. *Fire* **2023**, *6*, 280. [CrossRef]
11. Wotton, B.M.; Nock, C.A.; Flannigan, M.D. Forest fire occurrence and climate change in Canada. *Int. J. Wildland Fire* **2010**, *19*, 253–271. [CrossRef]
12. Ahn, Y.S.; Ryu, S.R.; Lim, J.H.; Lee, C.H.; Shin, J.H.; Choi, W.I.; Lee, B.D.; Jeong, J.-H.; An, K.W.; Seo, J.I. Effects of forest fires on forest ecosystems in eastern coastal areas of Korea and an overview of restoration projects. *Landsc. Ecol. Eng.* **2014**, *10*, 229–237. [CrossRef]
13. *The 6th Forest Master Plans*; Korea Forest Service: Daejeon, Republic of Korea, 2021.
14. Sakr, G.E.; Elhajj, I.H.; Mitri, G.; Wejinya, U.C. Artificial intelligence for forest fire prediction. In Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Montreal, QC, Canada, 6–9 July 2010; pp. 1311–1316. [CrossRef]
15. Chae, K.-J.; Lee, Y.R.; Cho, Y.J.; Park, J.H. Development of Gangwon Province Forest Fire Prediction Model using Machine Learning and Sampling. *Korea J. BigData* **2018**, *3*, 71–78. [CrossRef]
16. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [CrossRef]
17. Linardos, V.; Drakaki, M.; Tzionas, P.; Karnavas, Y.L. Machine Learning in Disaster Management: Recent Developments in Methods and Applications. *Mach. Learn. Knowl. Extr.* **2022**, *4*, 446–473. [CrossRef]
18. Shmuel, A.; Heifetz, E. Global Wildfire Susceptibility Mapping Based on Machine Learning Models. *Forests* **2022**, *13*, 1050. [CrossRef]
19. Jo, H.-W.; Krasovskiy, A.; Hong, M.; Corning, S.; Kim, W.; Kraxner, F.; Lee, W.-K. Modeling Historical and Future Forest Fires in South Korea: The FLAM Optimization Approach. *Remote Sens.* **2023**, *15*, 1446. [CrossRef]
20. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Aryal, J. Forest Fire Susceptibility and Risk Mapping Using Social/Infrastructural Vulnerability and Environmental Variables. *Fire* **2019**, *2*, 50. [CrossRef]
21. *Cadastral Statistical Annual Report*; Ministry of Land, Infrastructure and Transport: Sejong-si, Republic of Korea, 2022.
22. *Statistical Yearbook of Forestry*; Korea Forest Service: Daejeon, Republic of Korea, 2021.
23. *2019 Gangwon East Coast Forest Fire White Paper*; Ministry of Public Administration and Security: Sejong-si, Republic of Korea, 2019.
24. *Annual Climatological Report*; Korea Meteorological Administration: Seoul, Republic of Korea, 2022.
25. Hirose, K.; Kim, C.K.; Kim, C.S.; Chang, B.W.; Igarashi, Y.; Aoyama, M. Wet and dry deposition patterns of plutonium in Daejeon, Korea. *Sci. Total Environ.* **2009**, *332*, 243–252. [CrossRef]
26. Korea Meteorological Administration. Available online: http://climate.go.kr/home/ (accessed on 23 September 2023).
27. Daly, C.; Neilson, R.P.; Phillips, D.L. A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *J. Appl. Meteorol. Climatol.* **1994**, *33*, 140–158. [CrossRef]
28. Daly, C.; Gibson, W.; Taylor, G.; Johnson, G.; Pasteris, P. A knowledge-based approach to the statistical mapping of climate. *Clim. Res* **2002**, *22*, 99–113. [CrossRef]
29. Daly, C.; Helmer, E.H.; Quiñones, M. Mapping the climate of Puerto Rico, Vieques and Culebra. *Int. J. Clim* **2003**, *23*, 1359–1381. [CrossRef]
30. Kim, J.U.; Sang, J.; Kim, M.K.; Byun, Y.H.; Kim, D.H.; Kim, T.J. Future Climate Projection in South Korea using the High-Resolution SSP Scenarios based on Statistical Downscaling. *J. Clim. Res.* **2022**, *17*, 89–106. [CrossRef]
31. Lee, B.-D.; Koo, K.-S.; Lee, M.-B. Forest Fire Direction and Spread Characteristics by Field Investigations. *Fire Sci. Eng.* **2009**, *23*, 96–102.
32. Kim, S.J.; Lim, C.-H.; Kim, G.S.; Lee, J.; Geiger, T.; Rahmati, O.; Son, Y.; Lee, W.-K. Multi-Temporal Analysis of Forest Fire Probability Using Socio-Economic and Environmental Variables. *Remote Sens.* **2019**, *11*, 86. [CrossRef]
33. Danielsson, P.E. Euclidean distance mapping. *Comput. Graph. Image Process.* **1980**, *14*, 227–248. [CrossRef]
34. Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1900**, *50*, 157–175. [CrossRef]
35. Iban, M.C.; Sekertekin, A. Machine learning based wildfire susceptibility mapping using remotely sensed fire data and GIS: A case study of Adana and Mersin provinces, Turkey. *Ecol. Inform.* **2022**, *69*, 101647. [CrossRef]
36. Swets, J.A. Measuring the accuracy of diagnostic systems. *Science* **1988**, *240*, 1285–1293. [CrossRef]
37. Bradley, A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **1997**, *30*, 1145–1159. [CrossRef]

38. He, Q.; Jiang, Z.; Wang, M.; Liu, K. Landslide and Wildfire Susceptibility Assessment in Southeast Asia Using Ensemble Machine Learning Methods. *Remote Sens.* **2021**, *13*, 1572. [CrossRef]
39. Noce, S.; Collalti, A.; Santini, M. Likelihood of changes in forest species suitability, distribution, and diversity under future climate: The case of Southern Europe. *Ecol. Evol.* **2017**, *7*, 9358–9375. [CrossRef]
40. Khan, S.; Verma, S. Ensemble modeling to predict the impact of future climate change on the global distribution of *Olea europaea* subsp. *cuspidate*. *Front. For. Glob. Chang.* **2022**, *5*, 977691. [CrossRef]
41. Kim, J.S.; Park, S.Y. Landslide Susceptibility Mapping Using Ensemble FR and LR models at the Inje Area, Korea. *J. Korean Soc. Geospat. Inf. Sci.* **2017**, *25*, 19–27.
42. Ahn, H.G.; Lim, C.H. Assessing habitat suitability for timber species in South Korea under SSP scenarios. *Korean J. Environ. Biol.* **2022**, *40*, 567–578. [CrossRef]
43. Thuiller, W.; Lafourcade, B.; Engler, R.; Araújo, M.B. BIOMOD–A platform for ensemble forecasting of species distributions. *Ecography* **2009**, *32*, 369–373. [CrossRef]
44. Park, C.H. Deduction of Risk Assessment (RA) Items through the Characteristics and Vulnerabilities of Gangwon Fire. *J. Soc. Cult. Herit. Disaster Prev.* **2019**, *4*, 55–64.
45. Lee, C.J.; Lim, M.Y.; Lee, Y.H. Machine Learning for Big Data Analytics in Development of Wildfire Prediction Models. *J. Korean Soc. Hazard Mitig.* **2023**, *23*, 29–39. [CrossRef]
46. Rodrigues, M.; de la Riva, J. An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environ. Model. Softw.* **2014**, *57*, 192–201. [CrossRef]
47. Kalantar, B.; Ueda, N.; Idrees, M.O.; Janizadeh, S.; Ahmadi, K.; Shabani, F. Forest Fire Susceptibility Prediction Based on Machine Learning Models with Resampling Algorithms on Remote Sensing Data. *Remote Sens.* **2020**, *12*, 3682. [CrossRef]
48. Pang, Y.; Li, Y.; Feng, Z.; Feng, Z.; Zhao, Z.; Chen, S.; Zhang, H. Forest Fire Occurrence Prediction in China Based on Machine Learning Methods. *Remote Sens.* **2022**, *14*, 5546. [CrossRef]