

Article

Modeling and Mapping of Forest Fire Occurrence in the Lower Silesian Voivodeship of Poland Based on Machine Learning Methods

Slobodan Milanović ^{1,2,*} , Jan Kaczmarowski ³, Mariusz Ciesielski ⁴ , Zoran Trailović ⁵ , Miłosz Mielcarek ⁴ , Ryszard Szczygieł ⁶, Mirosław Kwiatkowski ⁶, Radomir Bałazy ⁷ , Michał Zasada ⁸  and Sladjan D. Milanović ⁹ 

¹ Department of Forestry, Faculty of Forestry, University of Belgrade, 11030 Belgrade, Serbia

² Department of Forest Protection and Wildlife Management, Faculty of Forestry and Wood Technology, Mendel University in Brno, 613 00 Brno, Czech Republic

³ General Directorate of the State Forests, 02-124 Warsaw, Poland

⁴ Department of Geomatics, Forest Research Institute, 05-090 Sękocin Stary, Poland

⁵ Department of Forest and Soil Sciences, Institute of Silviculture, University of Natural and Life Sciences, 1190 Vienna, Austria

⁶ Forest Fire Protection Laboratory, Forest Research Institute, 05-090 Sękocin Stary, Poland

⁷ Prevent Fires Foundation, 02-124 Warsaw, Poland

⁸ Department of Forest Management, Dendrometry and Forest Economics, Institute of Forest Sciences, Warsaw University of Life Sciences—SGGW, 02-124 Warsaw, Poland

⁹ Department for Biomechanics, Biomedical Engineering and Physics of Complex Systems, Institute for Medical Research, University of Belgrade, 11000 Belgrade, Serbia

* Correspondence: slobodan.milanovic@sfb.bg.ac.rs



Citation: Milanović, S.; Kaczmarowski, J.; Ciesielski, M.; Trailović, Z.; Mielcarek, M.; Szczygieł, R.; Kwiatkowski, M.; Bałazy, R.; Zasada, M.; Milanović, S.D.

Modeling and Mapping of Forest Fire Occurrence in the Lower Silesian Voivodeship of Poland Based on Machine Learning Methods. *Forests* **2023**, *14*, 46. <https://doi.org/10.3390/f14010046>

Academic Editors: Paul Sestras, Ștefan Bilașco, Mihai Nita and Sanda Roșca

Received: 21 November 2022

Revised: 21 December 2022

Accepted: 22 December 2022

Published: 26 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: In recent years, forest fires have become an important issue in Central Europe. To model the probability of the occurrence of forest fires in the Lower Silesian Voivodeship of Poland, historical fire data and several types of predictors were collected or generated, including topographic, vegetation, climatic, and anthropogenic features. The main objectives of this study were to determine the importance of the predictors of forest fire occurrence and to map the probability of forest fire occurrence. The H2O driverless artificial intelligence (DAI) cloud platform was used to model forest fire probability. The gradient boosted machine (GBM) and random forest (RF) methods were applied to assess the probability of forest fire occurrence. Evaluation the importance of the variables was performed using the H2O platform permutation method. The most important variables were the presence of coniferous forest and the distance to agricultural land according to the GBM and RF methods, respectively. Model validation was conducted using receiver operating characteristic (ROC) analysis. The areas under the curve (AUCs) of the ROC plots from the GBM and RF models were 83.3% and 81.3%, respectively. Based on the results obtained, the GBM model can be recommended for the mapping of forest fire occurrence in the study area.

Keywords: forest fire; ignition probability; random forest; gradient boosted machine

1. Introduction

At the global scale, forests are threatened by various abiotic and biotic agents. Among them, fire remains one of the most destructive factors with long-term impacts on forest ecosystems [1,2]. During the last two decades, fires have substantially impacted forest regions in Amazonia, North America, Siberia, Australia, and Africa [3]. In Europe, the Mediterranean region and the Balkan Peninsula have the highest incidences of forest fires and burnt areas [4]. In addition, Central European countries have experienced a substantial number of forest fires. Almost 8500 hectares of forest and forest land in Poland were completely decimated during the 2020 fire season [5]. During the fire season in 2022, extreme fire events occurred in the Bohemian Switzerland National Park, resulting in more

than 1400 hectares of forest in the Czech Republic being burnt (https://effis.jrc.ec.europa.eu/apps/effis_current_situation/index.html, accessed on 21 September 2022). At present, there is no safety zone for forest fires in Europe. Therefore, effective forest fire assessment systems are warranted for all parts of Europe, especially given the substantial costs of forest fire suppression [6]. This need was recognized in Poland in the early 1980s, resulting in the development of the first fire risk classification for state forests [7]. This was then followed by revision of the classification criteria in 1992 [8]. Currently, forest fire risk categories (KZPL) are determined according to the Regulation of the Minister of Environment of 2006, which details the principles of forest fire prevention (<https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20060580405>, accessed on 16 August 2022). Forest fire risk categories are revised during the development of forest management plans for specific forest districts every 10 years. This is undertaken by the Forest Research Institute, and the categories are determined based on the occurrence of fires; forest stand factors, including stand age, tree species, and forest habitat type; climatic conditions; and anthropogenic factors, such as population density. There are three categories of fire risk used in this context; namely, high, moderate, and low risk. The results from this categorization are used as the basis for all further planning of directional tasks in the field of fire protection. They also determine the needs in terms of technical infrastructure and the forest area that must be made available for firefighting operations. Forest fire occurrence using the Polish Forest Research Institute method is determined for selected forecast zones encompassing a group comprising several forest districts. The degree of risk is determined based on direct measurements of litter moisture, air temperature, and relative humidity at a height of 0.5 m above the ground, as well as the sum of precipitation within the preceding 24 h. Measurements were taken at 9:00 a.m. and 1:00 p.m. from 1 March to 30 September. Forest fire occurrence was predicted up to 24 h in advance [9]. The degree of risk for a given day is used to determine the level of readiness of the forest fire protection system. Similar variables describing fuel conditions, such as rainfall, temperature, wind speed, and relative humidity, have been used in other studies [10–12]. Vegetation indices have been widely used to predict fire risk [13–17] in forest fire occurrence modeling due to the strong influence of human activities on fire ignition [18–21]. Terrain-related factors, such as slope, elevation, and aspect, are usually derived from a digital elevation model (DEM). They are often included in the models for fire risk prediction, among other variables [22–25].

Most systems used in Europe for forest fire risk assessment—nationally [26] or at the continental level [27]—are based on meteorological data and are employed to determine whether the conditions for the occurrence of forest fires are present. Therefore, anthropogenic impacts are excluded from the risk assessment despite being the main cause of forest fires in Europe [27]. Anthropogenic impacts can be assessed through the effects of population density and specific behavior related to fire ignition, such as pastoral or agricultural burning, on the occurrence of forest fires. Other anthropogenic features, such as the proximity of an area to roads, railways, settlements, and agricultural land, can substantially contribute to forest fires. Therefore, anthropogenic features need to be incorporated into existing models used for forest fire risk assessment in addition to meteorological, topographic, and vegetation indices.

Various methods have been applied to evaluate the contributions of specific predictors to the modeling and mapping of the probability of the occurrence of forest fires. Logistic regression (LR) was the dominant method used during the 1990s and 2000s when dealing with binary outcomes, such as the presence or absence of fire [28–30]. Development of GIS technology allowed the wide use of the analytic hierarchy process (AHP) for forest fire risk mapping [31–33]. Later, machine learning (ML) methods developed in the context of artificial intelligence prevailed in wildfire science, with more than 300 articles being published on this topic since the 1990s [34]. Among the various ML methods, the random forest (RF) algorithm, which belongs to the decision tree branch of ML [35], has been demonstrated to be one of the most accurate [36–39]. However, in recent years, the gradient boosted machine (GBM) algorithm has emerged as a prominent method in forest fire

modeling, with almost equal performance to the RF algorithm [40,41]. The accuracy of these models varies from 60% to 86.7% [42–45] for LR through to 77.2% to 79.3% for the AHP [46–48], 82.5% to 98.9% for the RF model [45,49,50] and 84.2% to 88.9% in the GBM model [41,51].

The main objectives of this study were to: (i) determine the importance of the various meteorological, topographic, vegetation, and anthropogenic predictors for forest fire occurrence in the Lower Silesian Voivodeship of Poland using an approach based on machine learning methods; (ii) map the probability of forest fire occurrence for the study area based on developed models. To the best of our knowledge, this is the first study that has used machine learning methods to predict the probability of occurrence of forest fires in Poland by combining traditional and widely used variables.

2. Materials and Methods

2.1. Study Area

The Lower Silesian Voivodeship is located in the southwestern part of Poland on the border with the Czech Republic and Germany (Figure 1). Forest covers 29% of the study area. The forest of the Lower Silesian Voivodeship is managed by the Regional Directorate of State Forests (RDLP) in Wrocław, which is one of the 17 regional directorates of State Forests. The total area managed by the Wrocław RDLP is approximately 560,000 ha, of which 542,000 ha is forest area and the remaining area is non-forest land, such as agricultural land, wastelands, and water bodies. The spatial diversity of forest areas in the Wrocław RDLP area has led to the identification of two large, compact forest complexes. The first is located in the northwestern part of the RDLP, which is known as the Lower Silesian Forest, and the second southern part includes the forests in the Sudeten Mountains. Forests are found in small, fragmented complexes in the remaining areas (Figure 2b). The elevation ranges from 30 m asl up to 1603 m asl (Figure 2a) at the highest peak in the Sudeten Mountains [52]. The mean annual temperature ranges from 4 °C in the mountains to 9.4 °C in the lowlands, whilst the total annual precipitation ranges from 500 mm in the lowlands to 1300 mm in the mountains [53].

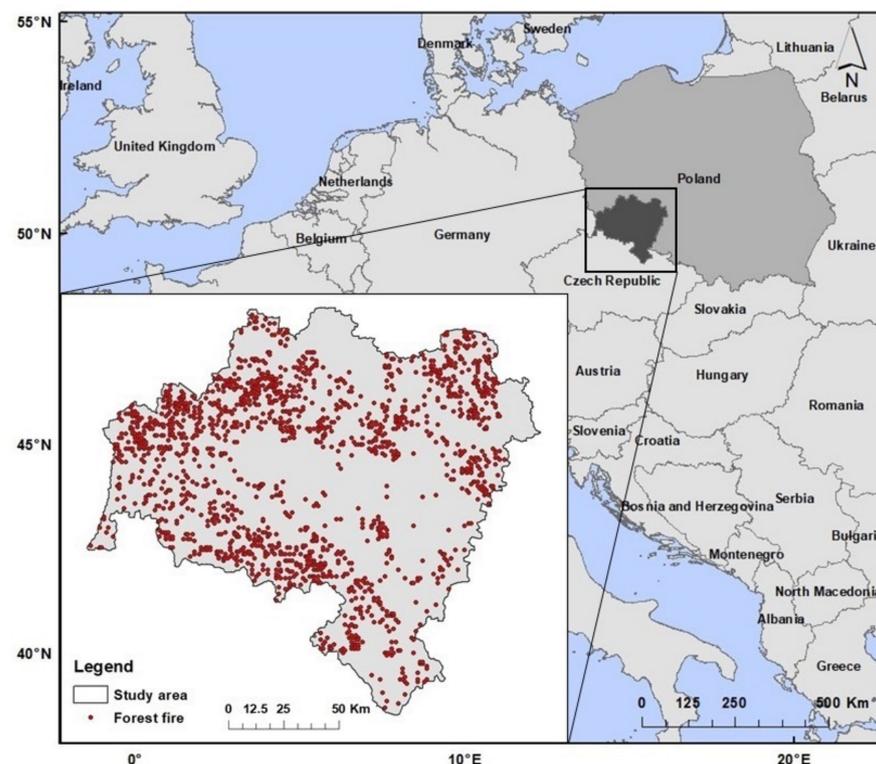


Figure 1. Study area: the Lower Silesian Voivodeship of Poland.

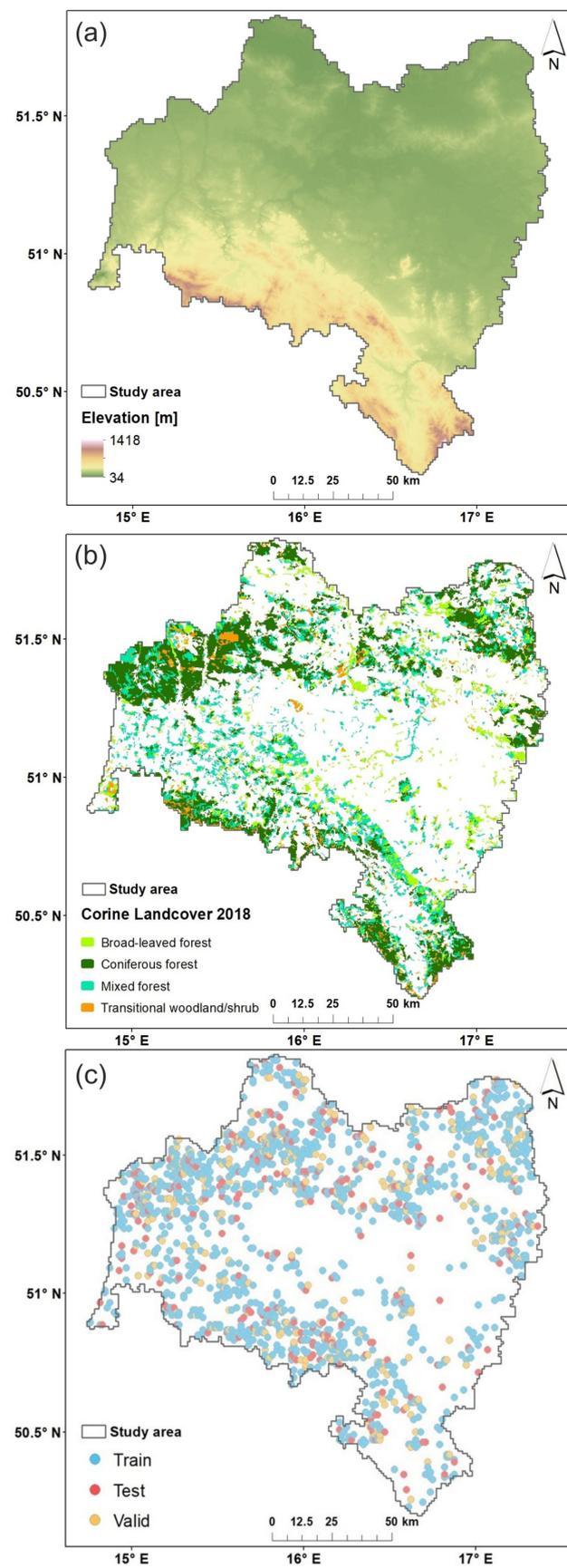


Figure 2. Altitude (a); selected CORINE land cover classes (b); and training, testing, and validation datasets (c) in the study area.

Owing to substantial geographical, soil, and climatic differences in the Wrocław RDLP area, there are distinct lowland (58%), upland (14%), and mountain (28%) habitats (Figure 2a). The forest structure is dominated by coniferous and mixed coniferous forests, which occupy 44% of the area, followed by mixed forests (35%) and deciduous forests (21%) (Figure 2b). Coniferous species constitute 72% of the area for all stands [54]. The main forest-forming species include pine (*Pinus* sp.) and larch (*Larix* sp.), comprising approximately 50% of all stands, as well as spruce (*Picea* sp.) (25%), oak (*Quercus* sp.), ash (*Fraxinus* sp.), maple (*Acer* sp.) (13%), birch (*Betula* sp.) and hornbeam (*Carpinus* sp.) (5%), beech (*Fagus sylvatica* L.) (5%), and alder (*Alnus* sp.) (3%) [54]. The age structure is dominated by stands aged 20–60 years [54].

2.2. Data Collection

2.2.1. Dependent Variable

Data on forest fire events were obtained from the Polish State Forest Information System (SILP) for the period from January 2010 to December 2020. All the fire events were identified using their coordinates during the period analyzed and were mapped to 1 × 1 km grid cells.

2.2.2. Independent Variables (Predictors)

Predictors were grouped into four main categories; namely, topography, vegetation, anthropogenic factors, and climate. Specific variables within the categories were selected based on previous studies on forest fire occurrence [28,45,55–61].

Topographic features, such as elevation, slope, aspect, the topographic wetness index (TWI), and solar radiation (SR), were derived from Shuttle Radar Topography Mission (SRTM) data with an average accuracy at the state level of precision of 6.2 m [62]. Average values for the elevation (E), dominant aspect (A), TWI, and SR were calculated for each polygon in a 1 × 1 km grid using ArcGIS software 10.2 (ESRI, Redlands, CA, USA).

Vegetation and land-cover data were obtained from the CORINE 2018 database (<https://land.copernicus.eu/pan-european/corine-land-cover/clc2018>, accessed on 20 December 2021). The vector layer was intersected with the polygon grid data. Objects in this vector layer were filtered for the following CORINE 2018 land-cover classes (CLC): broad-leaved forest (BF), coniferous forest (CF), mixed forest (MF), natural grassland (NG), moors and heathland (MH), transitional woodland–shrub (TWS), and sparsely vegetated areas (SVAs) (Figure 2b). By intersecting the polygon grid data with the polygon CLC layer filtered in this way, a new polygon layer with a table of attributes containing a polygon grid object ID, a CLC class ID with its description, and the area of the CLC class that fell into the respective grid polygon was generated.

Forest stand description data, such as the type of ground cover, habitat, and dominant tree species, were obtained from the Forest Data Bank (<https://www.bdl.lasy.gov.pl/portal/mapy-en>, accessed on 20 January 2022) and processed in the same way as the CORINE 2018 land-cover data.

Bioclimatic variables, precipitation during the fire season (PrecSe), and precipitation in the driest quarter (PrecDQ) were downloaded as GeoTiff files with a spatial resolution of 30 s (~1 km²) from the WorldClim portal (<https://www.worldclim.org/data/worldclim21.html>, accessed on 18 February 2022). Mean values for rainfall (Rain_13) and forest cover humidity (FCH_13), measured at 1 p.m., were calculated based on data collected from March to September (2010–2020) at permanent measurement locations across the study area. The period of the year (March to September) and the time span (2010–2020) correspond to the main fire season and the fire event dataset, respectively. The coordinates of the measurement locations and the values for rainfall (Rain_13) and forest cover humidity (FCH_13) for each day in the observed period were then converted into a spatial point layer with points representing measurement locations and a table of attributes with values for Rain_13 and FCH_13. This layer was then used for interpolation using the ordinary Kriging method [63], resulting in raster layers with pixels representing the values for rainfall (Rain_13) and the

forest cover humidity (FCH_13). Therefore, we were able to calculate the zonal statistics for the rainfall (Rain_13) and the forest cover humidity (FCH_13) for each polygon in the grid layer using the raster layers that represented the observed period.

Anthropogenic features, such as layers for roads, populated places, railroads, and agricultural land, were used to calculate the distance from each object in a 1×1 km grid to the nearest object on the respective layer. Therefore, the attribute table for the point and polygon grid layer generated was extended with the distances to the nearest building (DisBld), road (DisRo_B and DisRo_C), railway (DisRa), and agricultural land (DisAgL) (<https://www.openstreetmap.org>, accessed on 20 January 2022). The local road density (LRoD), regional road density (RRoD), and path density (PathD) in the grid were extracted from the same source and mapped to each cell in the grid. Socioeconomic data, such as the total number of inhabitants (TotIn), population density (PopD), population density in cities (PopD_C), unemployment rate (UR), mean salary (MS), tourist accommodation (TouAcc), and number of illegal landfills (NoIL), were extracted from the Statistics Poland Local Data Bank (<https://bdl.stat.gov.pl/bdl/start>, accessed on 20 December 2021) and mapped to each cell in the grid.

Following GIS analysis, Boolean values (yes/no) were assigned to the elements of the grid using the Spatial Join tool. “Yes” values were assigned to the elements where forest fires had historically occurred, and “no” values were assigned to those for which they had not. Each cell with at least one historical record of a fire event from 2010 to 2020 was classified as a fire cell and coded as “1” (one). In total, 1680 cells from the Lower Silesian Voivodeship were selected as fire cells and labeled with “1” for further analysis. The remaining 17,588 cells in the study area without fire events for the period reviewed were labeled with “0”.

After data acquisition, a database was created for the study area, and the next steps were undertaken as detailed in the workflow chart (Figure 3).

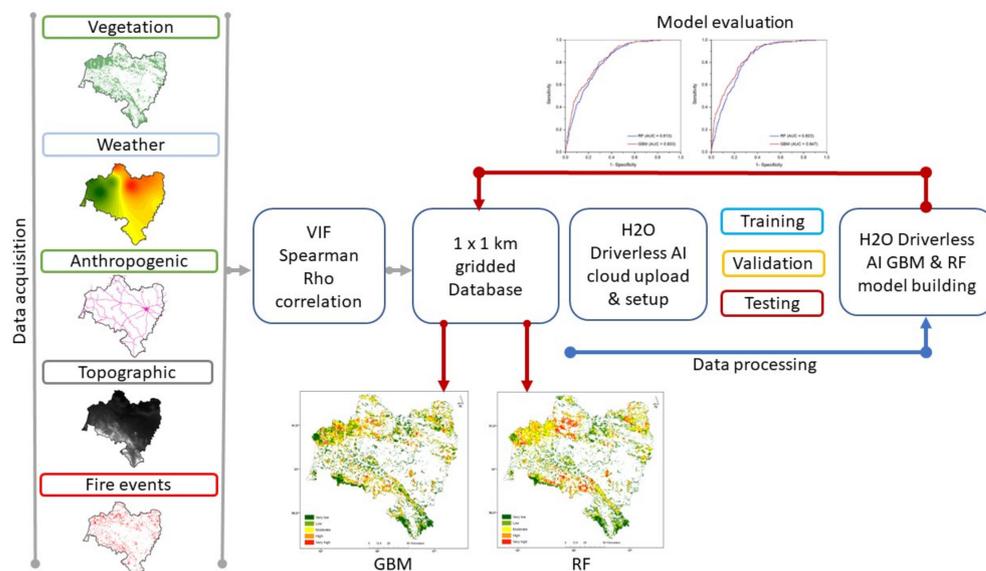


Figure 3. Data processing workflow.

2.3. Variable Selection

The multicollinearity of the variables was assessed in two parts. In the first part, explanatory variables were checked for multicollinearity using the variance inflation factor (VIF) and tolerance [64]. Only variables with a $VIF \leq 10$ and tolerance lower than 1 were considered for further analysis (Table A1). The second part was performed in four steps following the method proposed by Kuhn et al. [65]. Firstly, Spearman’s rho correlation matrix was calculated for all the remaining variables. Pairs of variables with a correlation coefficient higher than 0.7 were identified and removed in the second and third steps,

respectively. Finally, the correlation matrix was recalculated until no two variables had a correlation coefficient greater than 0.7 (Figure A1).

2.4. Model Training, Validation, and Testing

Prior to the creation of the models, data samples were randomly divided into training, validation, and test subsamples including 70%, 15%, and 15% of the fire and non-fire cells, respectively (Figure 2c). The H2O driverless artificial intelligence (DAI) platform was used to create both the GBM and RF models. This platform automates some of the most difficult machine-learning processes, such as feature engineering, model validation, model tuning, and model selection [66]. Given that the main DAI settings ranged from 1 to 10 in both models, the time, accuracy, and interpretability were set to 10. This setting allowed the platform to create a model without time limitations and with high accuracy and interpretability. It also made it possible to avoid the need for feature engineering. Models were created using a training dataset. The validation dataset was then used to tune the model pipeline, and the test dataset was used to test the model pipeline and create test predictions.

Model testing and validation were evaluated using receiver operating characteristic (ROC) analysis of the test and validation datasets. Area under curve (AUC) values between 0.5 and 0.7 indicated low precision, values between 0.7 and 0.8 indicated acceptable precision, values between 0.8 and 0.9 indicated excellent precision, and values higher than 0.9 indicated outstanding model precision [67].

Model evaluation was conducted using a confusion matrix for the entire dataset of the study area. For binary classification, a 2×2 matrix with rows representing the actual classes and columns indicating the predicted classes were used for model metrics. Based on this matrix, the following parameters were obtained: true positive (TP) indicated the number of positive samples correctly predicted as positive, true negative (TN) indicated the number of negative samples correctly predicted as negative, false positive (FP) indicated the number of negative samples incorrectly predicted as positive, and false negative (FN) indicated the number of positive samples incorrectly predicted as negative. From these four values, the accuracy (Acc) and precision (Prec) were calculated according to the following formulas:

$$\text{Acc} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN}) \quad (1)$$

$$\text{Prec} = (\text{TP}) / (\text{TP} + \text{FP}) \quad (2)$$

Finally, the predictive capacities of the models produced were compared using the distribution of fire cells in the study area across classes for the probability of forest fire occurrence. To designate each cell as fire or non-fire, the cutoff point was determined. For the GBM and RF models selected, the optimal cutoff point was determined using the sensitivity equal-specificity method [68] with the easyROC web tool [69]. The probability estimated for each cell was then compared to the optimal cutoff point. If the probability for one cell estimated by the obtained models was higher than the optimal cutoff point, that cell was then classified as a fire cell. In contrast, if the probability estimated was lower than the optimal cutoff point, a particular cell was then classified as a non-fire cell. Cutoff values were applied within the selected models to the dataset for the study area.

An additional model evaluation was conducted by comparing the fire event distribution among probability classes [45,60] and combinations of probability and vegetation classes to estimate the effect of land cover on ignition.

2.5. Evaluation of Variable Importance

Evaluation of the importance of the variables was performed using the H2O DAI platform permutation method. This explains the extent to which the performance of a model changes if the variable values are permuted. If the variable has low predictive power, shuffling its values should have a limited effect on model performance. However, if the variable is highly predictive, shuffling its values will reduce the model performance. The difference between the performance of the model before and after permutation of a

variable indicates the absolute importance of the permutation of that specific variable. For the calculation of the permutation importance, each variable was shuffled once.

2.6. Mapping the Probability of Forest Fire Occurrence

The probability of forest fire occurrence calculated by the GBM and RF models for fire cells and non-fire cells was used to map the probability of forest fire occurrence in the Wrocław RDLP area using ArcGIS 10.2. The produced map was classified into five categories—very low (0.01–0.40), low (0.21–0.65), medium (0.66–0.85), high (0.86–0.95), and very high (0.96–1.00) probabilities of forest fire occurrence—based on the percentile method proposed by Romero et al. [70].

3. Results

3.1. Contributions of Variables to Forest Fire Occurrence

From the 37 explanatory variables that met the conditions of having a VIF ≤ 10 and Spearman's rho correlation coefficient values lower than 0.7, 22 and 35 were included in the GBM and RF models, respectively (Table 1). The highest impact on the probability of fire in the GBM model had a relative contribution from coniferous forest (CF), as well as birch, aspen, poplar, willow, and goat willow as the dominant tree species (Spec_2); distance to agricultural land (DisAgL); leaf litter ground floor (Cov_1); and distance to roads (DisRo C). Regional road density (RRoD) had the lowest impact in the GBM model. In the RF model, the highest impact on the probability of fire was associated with distance to agricultural land (DisAgL), whilst the lowest impact on forest fire occurrence had relative contributions from heavily weedy ground floor (Cov_6) and transitional woodland–shrub (TWS) (Table 1).

3.2. Model Evaluation

The predictive capacity of the model was tested using classification ROC analysis. The AUCs of the ROC plots for the GBM and RF models were 84.7% and 82.3, respectively, for the testing data (Figure 4). For the validation data, the AUCs of the GBM and RF models were 83.3% and 81.3%, respectively.

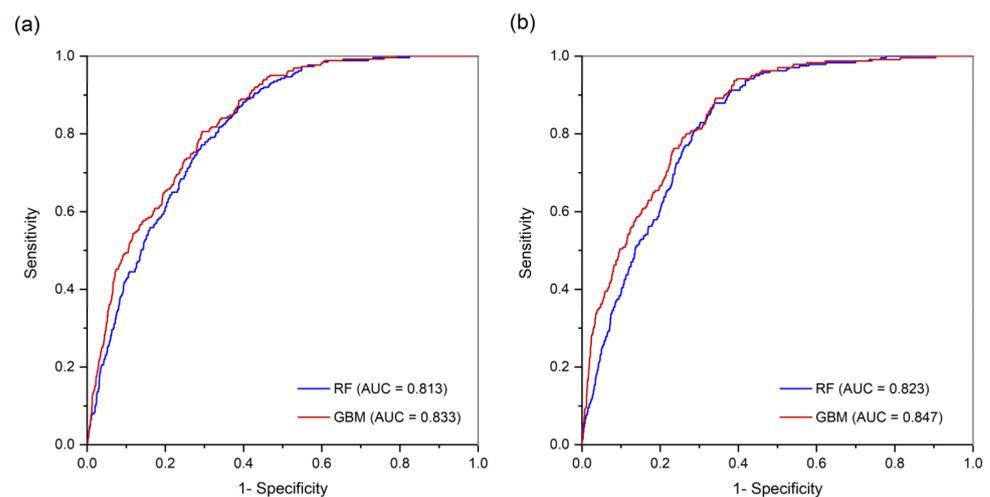


Figure 4. ROC analysis with GBM and RF models for validation (a) and test (b) datasets.

Table 1. Explanatory variable importance evaluation based on Gini impurity for RF and GBM models.

Variable Description	Code	Unit	GBM Relative Importance	RF Relative Importance
Vegetation				
Coniferous forest	CF	m ²	1.000	0.416
Mixed forest	MF	m ²		0.222
Transitional woodland–shrub	TWS	m ²		0.009
Leaf litter ground floor	Cov_1	m ²	0.627	0.325
Heavily turfed ground floor	Cov_5	m ²	0.388	0.866
Heavily weedy ground floor	Cov_6	m ²		0.007
Herbaceous ground floor	Cov_8	m ²		0.314
Fresh broadleaved and mixed broadleaved sites	Hab_4	m ²	0.297	0.331
Oaks, including red oak, common oak and pedunculate oak, black locust, bird cherry, hornbeam, sorb, ash, maple, sycamore, elm, linden, black walnut	Spec_2	m ²	0.448	0.222
Birch, aspen, poplar, willow, goat willow	Spec_4	m ²	0.711	0.104
Alder, gray alder	Spec_6	m ²		
Norway spruce	Spec_7	m ²	0.350	0.486
Anthropogenic				
Distance to Buildings	DisBld	m	0.478	0.022
Distance to Road B level	DisRo_B	m	0.494	0.032
Distance to road C level	DisRo_C	m	0.558	0.247
Distance to rail	DisRa	m	0.536	0.395
Distance to agricultural land	DisAgL	m	0.686	1.000
Total inhabitants	TotIn	N		0.037
Population density	PopD	N/km ²		0.251
Population density in cities	PopD_C	N/km ²		0.050
Unemployment rate	UR	%	0.347	0.278
Mean salary	MS	PLN		0.191
Tourists' accommodation	TouAcc	N/1000		0.317
Number of illegal landfills	NoIL	N/km ²		0.162
Local road density in grid	LRoD	km/km ²	0.380	0.161
Regional road density in grid	RRoD	km/km ²	0.244	0.063
Path density in grid	PathD	km/km ²		0.091
Topographic				
Distance to water	DisW	m		0.502
Aspect	A	Degrees		0.156
Elevation	E	m	0.421	0.332
Average solar radiation	A_SR	W/m ²	0.523	0.220
Topographic wetness index	TWI		0.425	0.111
Climatic				
Precipitation season	PrecSe		0.444	0.464
Precipitation in driest quarter	PrecDQ		0.383	0.237
Mean value of rainfall at 1 p.m. calculated on the basis of data from March to September (2010–2020)	Rain_13		0.412	0.345
Mean value of forest moisture humidity at 1 p.m. calculated on the basis of data from March to September (2010–2020)	FCH_13		0.343	0.202

3.3. Spatial Modeling of the Probability of Fire Occurrence

Zones with a very high probability for forest fire occurrence were situated in the northern and southwestern parts of the study area in both models and covered 6.4% (GBM)

and 7.4% (RF) of the study area. Zones with a very low probability for forest fire occurrence were situated in the central and southeastern parts of the study area and covered 29.1% (GBM) and 24.3% (RF) of the forested area (Figure 5).

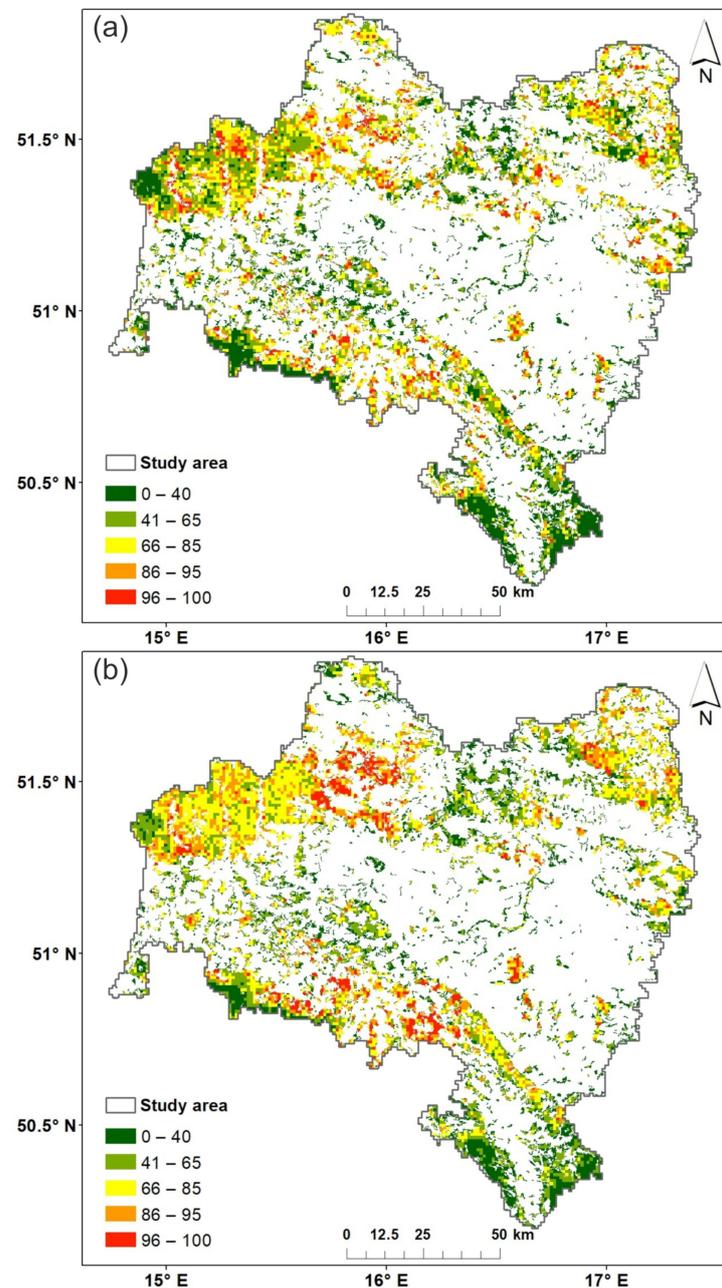


Figure 5. Maps of forest fire probability based on percentages with the (a) GBM model and (b) RF model.

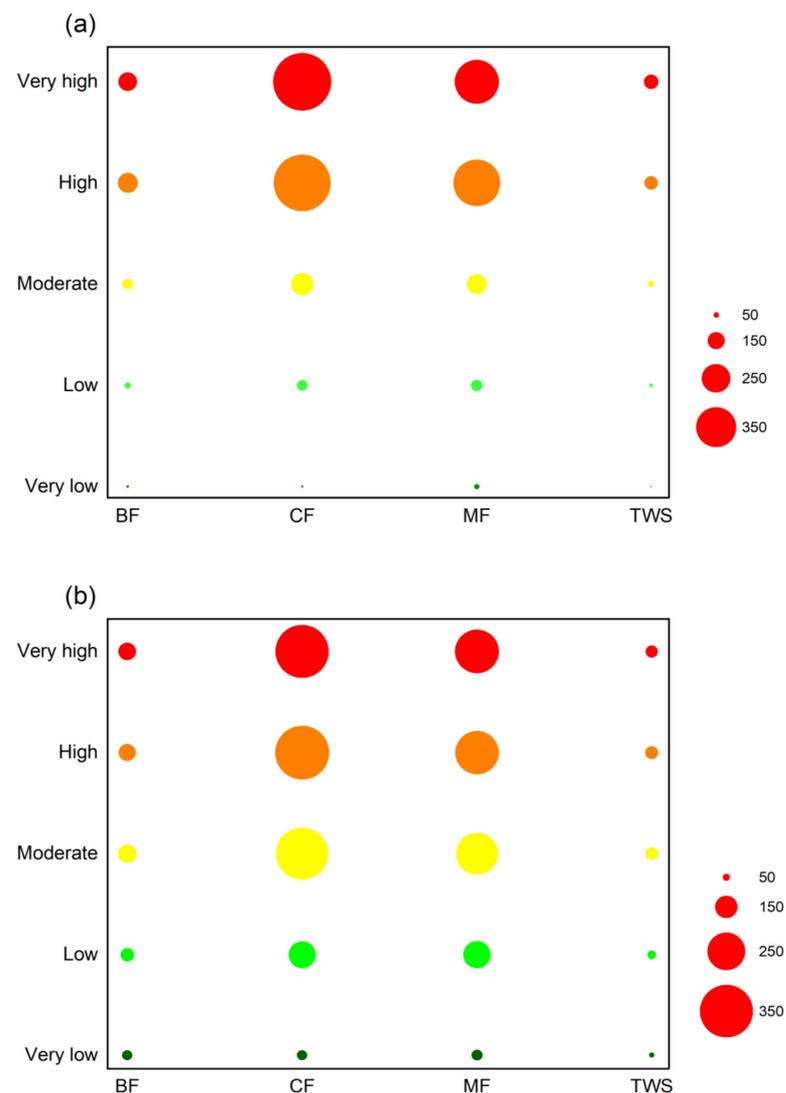
3.4. Model Validation

The GBM model performed better for the very low, low, and moderate risk classes than the RF model in terms of identifying lower forest fire incidences. The GBM model performed better for the high and very high risk classes, identifying higher incidences of forest fires than the RF model (Table 2). When the aggregated incidence for the three highest classes (moderate, high, and very high) was compared between the models, there was a very slight difference between the GBM and RF models (93% vs. 86%). In contrast, the GBM model classified only 6.7% of forest fire events in the two lowest classes (low and very low), whereas the RF model classified 13.7% of forest fire events in the same group of classes in the study area.

Table 2. Forest fire distribution across different probability classes based on GBM and RF models for the study area.

Forest Fire Probability Percentile	Forest Fire Probability Class	GBM	RF
0–40	Very low	13	35
41–65	Low	100	196
66–85	Moderate	224	460
86–95	High	550	487
96–100	Very high	793	502

Both models highlighted the contribution of coniferous forests to fire incidence, whilst only the RF model showed moderate importance for mixed forests and the marginal effect of transitional woodland and shrubs on fire incidence (Figure 6). In the GBM model, the incidence of forest fire was concentrated in the high and very high risk zones with fire-prone vegetation types, while in the RF model, the fire incidence spread into the moderate risk zone.

**Figure 6.** Forest fire distribution across vegetation and occurrence probability classes defined by the (a) GBM and (b) RF models.

The predictive capacity of the models was assessed using classification tables comparing the observed and predicted values for the study area dataset after model tuning

with the cutoff method. According to the overall classification, the GBM and RF models correctly classified 85.2% and 77.8% of cases, respectively, when the models were applied to the entire dataset (Table 3). The AUCs of the ROC plots for the GBM and RF models were 93.7% and 86.8%, respectively. The ROC analysis demonstrated the high levels of precision of the GBM and RF models, respectively (Figure 7).

Table 3. Classification tables for the study area dataset based on GBM and RF models after application of cutoff values based on the sensitivity equal-specificity method, along with accuracy and precision values.

Model	Cutoff	Observed	Predicted		Acc (%)	Prec (%)
			0	1		
GBM	0.1326	Observed	0	14,987	85.22	85.25
			1	248		
RF	0.1357	Observed	0	13,685	77.81	77.81
			1	373		

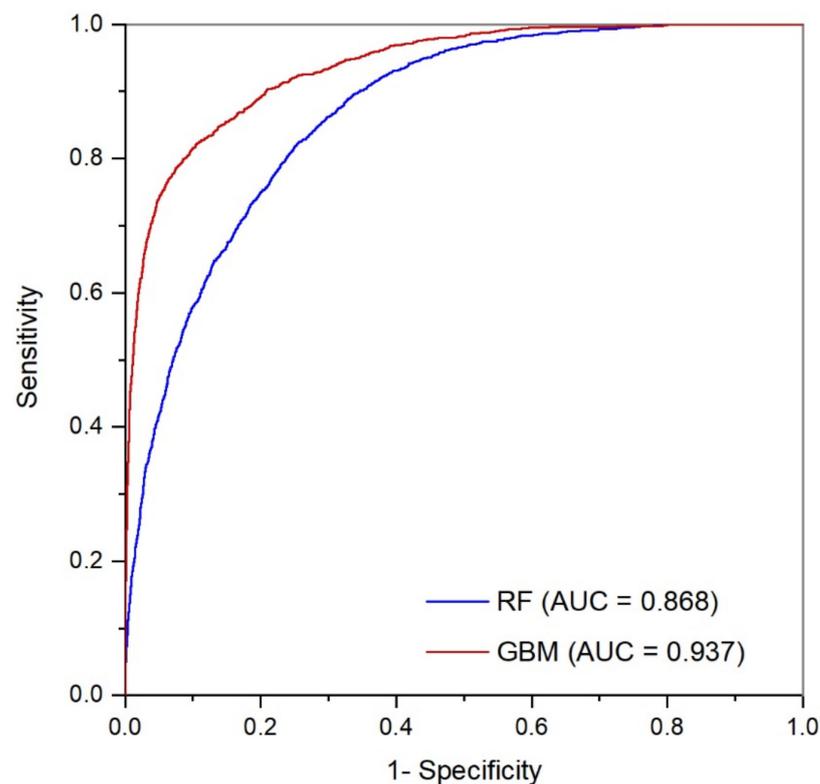


Figure 7. ROC analyses of the GBM and RF models for all fire data in the study area.

4. Discussion

Assessing the probability of the occurrence of forest fires and its drivers is of the highest importance for all endangered regions. In this study, the impacts of vegetation, anthropogenic, topographic, and climatic factors on forest fire ignition were evaluated using a ML approach. The GBM model used 22 variables to predict the probability of the occurrence of forest fires, making it more applicable for smaller datasets [71,72] compared to the RF model, which included 35 variables. Vegetation features, such as coniferous forest (CF), were found to be the most significant factor contributing to the probability of the occurrence of forest fires in the GBM model. This can be explained by the dominance of coniferous (mostly pine) forests in the study area [54], which are highly prone to forest

fires [73,74]. Anthropogenic features, such as distance to agricultural land (DisAgL), were the second most important contributing variable in the GBM model, but they were the most important according to the RF model. Such a high ranking for this variable in both models is consistent with the findings of a previous study [45] that highlighted the wildland–urban interface as the main driving factor for forest fire ignition [75–77]. The importance given by the GBM model to other anthropogenic features, such as distances to buildings, roads, and railways, confirmed the role of the wildland–urban interface (WUI) in forest fire ignition. The distance to railways had a higher level of importance than the distance to roads in the GBM model. In the RF model, the opposite relationship was recorded. Railways have often been related to accidental or negligent fires. Together with road density, railways strongly affect the probability of fires by increasing the numbers of residents and visitors [20]. A similar study in China showed a positive effect from railway density on fire ignition, while road density had a negative impact on fire ignition [21]. In contrast, research by Pinto et al. [19] found a positive effect from road density on fire incidence.

Many of the variables related to ground cover type, habitat, and species composition—part of an old but highly effective Polish system for fire risk assessment—showed very high importance in both models. It was highly important to test the ability to implement those variables in the new methods for fire risk assessment and to compare their contributions to other widely used variables. However, the reproducibility of the models that use these variables is limited to Poland due to lack of such data in other countries. Forest floor cover types had a relatively high impact on forest fire in both models and were ranked third (e.g., leaf litter ground floor (Cov_1)) and second (e.g., heavily turfed ground floor (Cov_5)) in the GBM and RF models, respectively. Forest floor cover is considered a fine fuel [78] and is, therefore, strongly influenced by weather conditions [79–81]. It is important for ignition of initial forest fires [73,82]. Initial fires usually start on the forest floor and may develop into ground-type forest fires or, eventually, into crown fires. However, this only occurs if the weather conditions are favorable for the burning process. In contrast, if the forest floor cover is unsuitable for burning, the initial forest fire will end and there will be no forest fire [83]. The importance of the weather conditions for fire ignition, which is emphasized in many studies [81,83–85], was also confirmed in our study, particularly by the GBM model. Topographic variables showed a relatively high contribution to the probability of the occurrence of forest fires in our study. Solar radiation, the topographic wetness index, and elevation were included in the GBM model as variables, while aspect and distance to water were dropped by the algorithm. In contrast, these two variables showed the highest contributions to the probability of the occurrence of forest fires among the topographic features in the RF model. In general, topographic features modify the fuel and its ability to burn [86–90].

Both models showed excellent performance in the validation and test datasets. However, the GBM model performed slightly better than the RF model based on the AUC values. The relationship observed between the performances of the GBM and RF models in prediction of fire occurrence was consistent with similar studies in other regions [40,51]. The GBM model shifted its performance from being excellent in the validation and test datasets to outstanding when the obtained model was applied to the study area dataset. Meanwhile, the RF model maintained the same performance when the model was applied to the study area dataset as for the validation and test datasets.

The zones with the highest probability for forest fire ignition were located in the central and northern parts of the study area in both models. The lowest probability was found for the south and southwestern parts. The dominance of mountains with climatic conditions less favorable for burning [91] explains the lower probability for the occurrence of forest fires in the south and southwestern areas. In contrast, the flat areas in the central and northern parts of the study area exhibit more pronounced drought periods during summer, explaining the higher fire incidence. Higher temperatures and a lack of precipitation lead to lower fuel moisture content [92], which makes the fuel and the area more susceptible to ignition [82,93–95]. Low precipitation has been described

as a determining factor for ignition [96]. Mapped zones with higher forest fire ignition probabilities in the central and northern parts of the study area need to be intensively monitored during the fire season. Such a measure would allow early detection of fires and rapid response to them once detected, leading to a reduction in the burnt area. Fire-smart silvicultural measures [97] can be applied in zones with higher probabilities of ignition to reduce fire risk. Therefore, in addition to growing less flammable tree species [74], which should be mandatory for afforestation, silvicultural measures, such as pruning lower branches, need to be implemented to prevent the transfer of surface fire to crowns in vulnerable zones. At the landscape level, fuel reduction treatments, such as thinning, prescribed burning, and fuel breaks [78,98,99], can be incorporated into forest management plans as effective tools to mitigate forest fire risk.

Ciesielski et al. [61] used logistic regression to evaluate the contributions of variables to forest fire occurrence with weather features excluded as predictors. Kolenek et al. [100] examined the impact of human activity on the occurrence of fires, emphasizing them to the detriment of other groups of predictors. In this study, we explored the potential for the incorporation of traditional variables used in Poland alongside widely used variables in the prediction of forest fire occurrence using an innovative machine learning technique. The produced maps can be used by forest professionals for the planning of long-term measures aiming to decrease fire risk, especially in very vulnerable protected areas [101]. Some structural and non-structural measures that can reduce the risks related to arson can only be implemented in urban areas if they are situated near forests, which is one of the limitations of this study focused only on forest fires. Further studies need to focus more on the assessment of both forested and urban areas to complete the picture of fire risk in the Lower Silesian Voivodeship of Poland.

5. Conclusions

Regions considered to be resistant to extreme fire events will require effective systems protecting them against forest fires in the future. The implementation of a machine learning approach in fire risk assessment may be a potential solution to this problem. The GBM model identified the coniferous forests, the distance to agricultural land, and leaf litter as the most important driving factors for the occurrence of fires. The RF model identified the distance to agricultural land, coniferous forests, and heavily turfed ground as the most important predictors for the occurrence of forest fires. Both models displayed relatively high predictive abilities. However, the GBM model was more efficient and may be recommended for mapping the occurrences of forest fires in the Lower Silesian Voivodeship of Poland.

Author Contributions: Conceptualization, S.M., J.K. and M.C.; methodology, S.M., J.K. and M.C.; software, J.K. and Z.T.; validation, M.Z., M.C., J.K. and S.D.M.; formal analysis, J.K. and Z.T.; investigation, S.M., M.C., J.K. and Z.T.; resources, M.K.; data curation, M.C., M.M., M.K. and R.B.; writing—original draft preparation, S.M.; writing—review and editing, M.C. and S.D.M.; visualization, Z.T.; supervision, R.S.; project administration, R.S. and M.K.; funding acquisition, R.S. and S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Education, Science, and Technological Development, Republic of Serbia, grant numbers 451-03-68/2022-14/200169 and 451-03-68/2022-14/200015 and co-financed by the Polish State Forests, grant numbers 500 477 and 500 446.

Data Availability Statement: The data presented in this study are available upon reasonable request from the corresponding author.

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

Appendix A

Table A1. Independent variables considered for forest fire occurrence models with codes, units, sources, tolerance and VIF.

Variable	Code	Unit	Source	Tolerance	VIF
Vegetation					
Broad-leaved forest	BF	m ²	CORINE 2018	0.437	2.287
Coniferous forest	CF	m ²		0.407	2.458
Mixed forest	MF	m ²		0.468	2.137
Transitional woodland–shrub	TWS	m ²		0.779	1.284
Leaf litter ground floor	Cov_1	m ²	Polish State Forest Information System (SILP)	0.629	1.589
Heavily turfed ground floor	Cov_5	m ²		0.568	1.761
Heavily weedy ground floor	Cov_6	m ²		0.645	1.551
Herbaceous ground floor	Cov_8	m ²		0.632	1.582
Fresh broadleaved and mixed broadleaved sites	Hab_4	m ²		0.713	1.402
Oaks, including red oak, common oak and pedunculate oak, black locust, bird cherry, hornbeam, sorb, ash, maple, sycamore, elm, linden, black walnut	Spec_2	m ²		0.485	2.062
Birch, aspen, poplar, willow, goat willow	Spec_4	m ²		0.688	1.454
Alder, gray alder	Spec_6	m ²		0.715	1.398
Norway spruce	Spec_7	m ²	0.277	3.605	
Anthropogenic					
Distance to buildings	DisBld	m	OpenStreetMap	0.434	2.306
Distance to road B level	DisRo_B	m		0.742	1.349
Distance to road C level	DisRo_C	m		0.625	1.601
Distance to rail	DisRa	m		0.864	1.157
Distance to agricultural land	DisAgL	m	CORINE 2018	0.442	2.263
Total inhabitants	TotIn	N		0.419	2.385
Population density	PopD	N/km ²	Statistics PolandDB	0.189	5.287
Population density in cities	PopD_C	N/km ²		0.232	4.307
Unemployment rate	UR	%		0.459	2.179
Mean salary	MS	PLN		0.734	1.362
Tourists' accommodation	TouAcc	N/1000		0.624	1.602
Number of illegal landfills	NoIL	N/km ²		0.484	2.065
Local road density in grid	LRoD	km/km ²		0.524	1.907
Regional road density in grid	RRoD	km/km ²		0.656	1.525
Path density in grid	PathD	km/km ²	0.478	2.092	
Topographic					
Distance to water	DisW	m	OpenStreetMap	0.672	1.488
Aspect	A	Degrees		0.895	1.117
Elevation	E	m	DEM	0.183	5.45
Average solar radiation	A_SR	W/m ²		0.779	1.284
Topographic wetness index	TWI			0.899	1.112
Climatic					
Precipitation during fire season	PrecSe		WorldClim	0.463	2.159
Precipitation in driest quarter	PrecDQ			0.531	1.884
Mean value of rainfall at 1 p.m. calculated on the basis of data from March to September (2010–2020)	Rain_13		Meteorological stations/measurement points (State Forests)	0.217	4.603
Mean value of forest moisture humidity at 1 p.m. calculated on the basis of data from March to September (2010–2020)	FCH_13			0.236	4.229

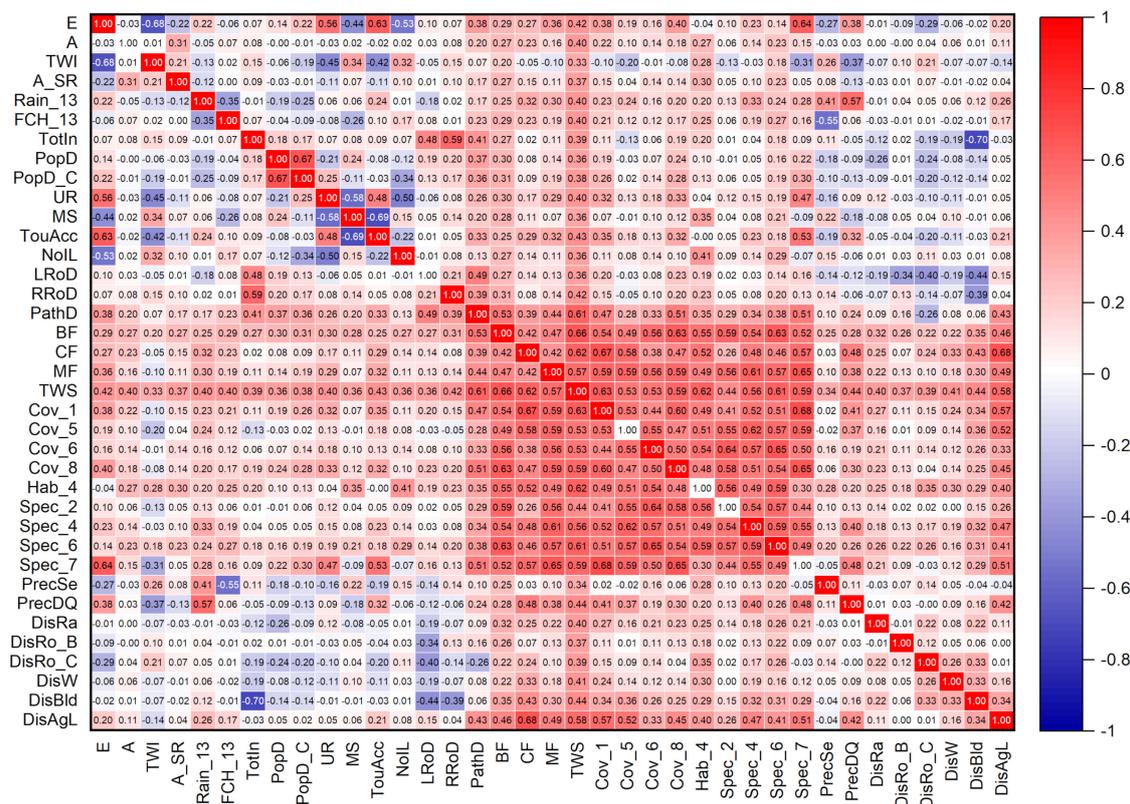


Figure A1. Correlation plot for all preselected variables based on Spearman’s rho coefficient.

References

1. Kastridis, A.; Stathis, D.; Sapountzis, M.; Theodosiou, G. Insect Outbreak and Long-Term Post-fire Effects on Soil Erosion in Mediterranean Suburban Forest. *Land* **2022**, *11*, 911. [\[CrossRef\]](#)
2. Nelson, A.R.; Narrowe, A.B.; Rhoades, C.C.; Fegell, T.S.; Daly, R.A.; Roth, H.K.; Chu, R.K.; Amundson, K.K.; Young, R.B.; Steindorff, A.S.; et al. Wildfire-Dependent Changes in Soil Microbiome Diversity and Function. *Nat. Microbiol.* **2022**, *7*, 1419–1430. [\[CrossRef\]](#)
3. Pereira, J.M.C.; Oom, D.; Silva, P.C.; Benali, A. Wild, Tamed, and Domesticated: Three Fire Macroregimes for Global Pyrogeography in the Anthropocene. *Ecol. Appl.* **2022**, *32*, e2588. [\[CrossRef\]](#)
4. Senf, C.; Seidl, R. Storm and Fire Disturbances in Europe: Distribution and Trends. *Glob. Chang. Biol.* **2021**, *27*, 3605–3619. [\[CrossRef\]](#) [\[PubMed\]](#)
5. San-Miguel-Ayanz, J.; Durrant, T.; Boca, R.; Maianti, P.; Liberta', P.; Artes Vivancos, T.; Jacome Felix Oom, D.P.; Branco, A.; de Rigo, D.; Ferrari, D.; et al. *Forest Fires in Europe, Middle East and North Africa 2020*; European Commission’s Joint Research Centre: Ispra, Italy, 2021; Volume 174. [\[CrossRef\]](#)
6. Mattioli, W.; Ferrara, C.; Lombardo, E.; Barbati, A.; Salvati, L.; Tomao, A. Estimating Wildfire Suppression Costs: A Systematic Review. *Int. For. Rev.* **2022**, *24*, 15–29. [\[CrossRef\]](#)
7. Karlikowski, T.; Łonkiewicz, B. *Badania i Ustalenie Kryteriów Oceny Zagrożenia Pożarowego Lasu w Oparciu o Warunki Meteorologiczne i Skład Gatunkowy Drzewostanu*; Raport IBL; IBL: Warszawa, Poland, 1975.
8. Santorski, Z. Regionalization of Forest Fire Danger in Poland. Ph.D. Thesis, Forest Research Institute, Warsaw, Poland, 1994.
9. Szczygieł, R.; Kwiatkowski, M.; Kołakowski, B.; Piwnicki, J. Dynamic Forest Fire Risk Evaluation in Poland. *Folia For. Pol. A* **2020**, *62*, 139–144. [\[CrossRef\]](#)
10. Vasilakos, C.; Kalabokidis, K.; Hatzopoulos, J.; Matsinos, I. Identifying Wildland Fire Ignition Factors through Sensitivity Analysis of a Neural Network. *Nat. Hazards* **2009**, *50*, 125–143. [\[CrossRef\]](#)
11. Cortez, P.; Morais, A.; Data Mining, A. Approach to Predict Forest Fires Using Meteorological Data. New Trends in Artificial Intelligence. In Proceedings of the 13th EPIA 2007–Portuguese Conference on Artificial Intelligence, Guimarães, Portugal, 3–7 December 2017; pp. 512–523.
12. Little, J.K.; Prior, L.D.; Williamson, G.J.; Williams, S.E.; Bowman, D.M.J.S. Fire Weather Risk Differs across Rain Forest–Savanna Boundaries in the Humid Tropics of North-Eastern Australia: Fire Weather Across a Rain Forest Boundary. *Austral Ecol.* **2012**, *37*, 915–925. [\[CrossRef\]](#)
13. Singh, M.; Huang, Z. Analysis of Forest Fire Dynamics, Distribution and Main Drivers in the Atlantic Forest. *Sustainability* **2022**, *14*, 992. [\[CrossRef\]](#)

14. Mohammadpour, P.; Viegas, D.X.; Viegas, C. Vegetation Mapping with Random Forest Using Sentinel 2 and GLCM Texture Feature—A Case Study for Lousã Region, Portugal. *Remote Sens.* **2022**, *14*, 4585. [[CrossRef](#)]
15. Rabiei, J.; Khademi, M.S.; Bagherpour, S.; Ebadi, N.; Karimi, A.; Ostad-Ali-Askari, K. Investigation of Fire Risk Zones Using Heat–Humidity Time Series Data and Vegetation. *Appl. Water Sci.* **2022**, *12*, 216. [[CrossRef](#)]
16. Li, R.; Fu, Y.; Bergeron, Y.; Valeria, O.; Chavardès, R.D.; Hu, J.; Wang, Y.; Duan, J.; Li, D.; Cheng, Y. Assessing Forest Fire Properties in Northeastern Asia and Southern China with Satellite Microwave Emissivity Difference Vegetation Index (EDVI). *ISPRS J. Photogramm. Remote Sens.* **2022**, *183*, 54–65. [[CrossRef](#)]
17. Rodrigues Silva, F.G.; Santos, A.R.; Fiedler, N.C.; Paes, J.B.; Alexandre, R.S.; Guerra Filho, P.A.; Silva, R.G.; Moura, M.M.; Silva, E.F.; Silva, S.F.; et al. Geotechnology Applied to Analysis of Vegetation Dynamics and Occurrence of Forest Fires on Indigenous Lands in Cerrado-Amazonia Ecotone. *Sustainability* **2022**, *14*, 6919. [[CrossRef](#)]
18. Guo, F.; Su, Z.; Wang, G.; Sun, L.; Lin, F.; Liu, A. Wildfire Ignition in the Forests of Southeast China: Identifying Drivers and Spatial Distribution to Predict Wildfire Likelihood. *Appl. Geogr.* **2016**, *66*, 12–21. [[CrossRef](#)]
19. Pinto, G.A.S.J.; Rousseu, F.; Niklasson, M.; Drobyshev, I. Effects of Human-Related and Biotic Landscape Features on the Occurrence and Size of Modern Forest Fires in Sweden. *Agric. For. Meteorol.* **2020**, *291*, 108084. [[CrossRef](#)]
20. Arndt, N.; Vacik, H.; Koch, V.; Arpaci, A.; Gossow, H. Modeling Human-Caused Forest Fire Ignition for Assessing Forest Fire Danger in Austria. *iForest* **2013**, *6*, 315–325. [[CrossRef](#)]
21. Su, Z.; Tigabu, M.; Cao, Q.; Wang, G.; Hu, H.; Guo, F. Comparative Analysis of Spatial Variation in Forest Fire Drivers between Boreal and Subtropical Ecosystems in China. *For. Ecol. Manag.* **2019**, *454*, 117669. [[CrossRef](#)]
22. Nasiri, V.; Sadeghi, S.M.M.; Bagherabadi, R.; Moradi, F.; Deljouei, A.; Borz, S.A. Modeling Wildfire Risk in Western Iran Based on the Integration of AHP and GIS. *Environ. Monit. Assess.* **2022**, *194*, 644. [[CrossRef](#)]
23. Si, L.; Shu, L.; Wang, M.; Zhao, F.; Chen, F.; Li, W.; Li, W. Study on Forest Fire Danger Prediction in Plateau Mountainous Forest Area. *Nat. Hazards Res.* **2022**, *2*, 25–32. [[CrossRef](#)]
24. Dhar, T.; Bhatta, B.; Aravindan, S. Forest Fire Occurrence, Distribution and Risk Mapping Using Geoinformation Technology: A Case Study in the Sub-tropical Forest of the Meghalaya, India. *Remote Sens. Appl.* **2023**, *29*, 100883. [[CrossRef](#)]
25. Li, W.; Xu, Q.; Yi, J.; Liu, J. Predictive Model of Spatial Scale of Forest Fire Driving Factors: A Case Study of Yunnan Province, China. *Sci. Rep.* **2022**, *12*, 19029. [[CrossRef](#)] [[PubMed](#)]
26. Camia, A.; Amatulli, G. Weather Factors and Fire Danger in the Mediterranean. In *Earth Observation of Wildland Fires in Mediterranean Ecosystems*; Chuvieco, E., Ed.; Springer: Berlin/Heidelberg, Germany, 2009. [[CrossRef](#)]
27. Fujioka, F.M.; Gill, A.M.; Viegas, D.X.; Wotton, B.M. Chapter 21. Fire Danger and Fire Behavior Modeling Systems in Australia, Europe, and North America. In *Developments in Environmental Science*; Elsevier: Amsterdam, The Netherlands, 2008; Volume 8, pp. 471–497. [[CrossRef](#)]
28. Catry, F.X.; Rego, F.C.; Bação, F.L.; Moreira, F. Modeling and Mapping Wildfire Ignition Risk in Portugal. *Int. J. Wildland Fire* **2009**, *18*, 921–931. [[CrossRef](#)]
29. Andrews, P.L.; Loftsgaarden, D.O.; Bradshaw, L.S. Evaluation of Fire Danger Rating Indexes Using Logistic Regression and Percentile Analysis. *Int. J. Wildland Fire* **2003**, *12*, 213–226. [[CrossRef](#)]
30. Zhang, Z.X.; Zhang, H.Y.; Zhou, D.W. Using GIS Spatial Analysis and Logistic Regression to Predict the Probabilities of Human-Caused Grassland Fires. *J. Arid Environ.* **2010**, *74*, 386–393. [[CrossRef](#)]
31. Nuthammachot, N.; Stratoulas, D. A GIS- and AHP-Based Approach to Map Fire Risk: A Case Study of Kuan Kreng Peat Swamp Forest, Thailand. *Geocarto Int.* **2021**, *36*, 212–225. [[CrossRef](#)]
32. Novo, A.; Fariñas-Álvarez, N.; Martínez-Sánchez, J.; González-Jorge, H.; Fernández-Alonso, J.M.; Lorenzo, H. Mapping Forest Fire Risk—A Case Study in Galicia (Spain). *Remote Sens.* **2020**, *12*, 3705. [[CrossRef](#)]
33. Eskandari, S.; Miesel, J.R. Comparison of the Fuzzy AHP Method, the Spatial Correlation Method, and the Dong Model to Predict the Fire High-Risk Areas in Hyrcanian Forests of Iran. *Geom. Nat. Hazards Risk* **2017**, *8*, 933–949. [[CrossRef](#)]
34. Jain, P.; Coogan, S.C.P.; Subramanian, S.G.; Crowley, M.; Taylor, S.; Flannigan, M.D. A Review of Machine Learning Applications in Wildfire Science and Management. *Environ. Rev.* **2020**, *28*, 478–505. [[CrossRef](#)]
35. Breiman, L. Statistical Modeling: The Two Cultures. *Stat. Sci.* **2001**, *16*, 199–215. [[CrossRef](#)]
36. Ghorbanzadeh, O.; Valizadeh Kamran, K.V.; Blaschke, T.; Aryal, J.; Naboureh, A.; Einali, J.; Bian, J. Spatial Prediction of Wildfire Susceptibility Using Field Survey GPS Data and Machine Learning Approaches. *Fire* **2019**, *2*, 43. [[CrossRef](#)]
37. Gigović, L.; Pourghasemi, H.R.; Drobnyak, S.; Bai, S. Testing a New Ensemble Model Based on SVM and Random Forest in Forest Fire Susceptibility Assessment and Its Mapping in Serbia’s Tara National Park. *Forests* **2019**, *10*, 408. [[CrossRef](#)]
38. Janiec, P.; Gadal, S. A Comparison of Two Machine Learning Classification Methods for Remote Sensing Predictive Modeling of the Forest Fire in the North-Eastern Siberia. *Remote Sens.* **2020**, *12*, 4157. [[CrossRef](#)]
39. Xu, Z.; Liu, D.; Yan, L. Temperature-Based Fire Frequency Analysis Using Machine Learning: A Case of Changsha, China. *Clim. Risk Manag.* **2021**, *31*, 100276. [[CrossRef](#)]
40. Achu, A.L.; Thomas, J.; Aju, C.D.; Gopinath, G.; Kumar, S.; Reghunath, R. Machine-Learning Modelling of Fire Susceptibility in a Forest-Agriculture Mosaic Landscape of Southern India. *Ecol. Inform.* **2021**, *64*, 101348. [[CrossRef](#)]
41. Banerjee, P. MODIS-FIRMS and Ground-Truthing-Based Wildfire Likelihood Mapping of Sikkim Himalaya Using Machine Learning Algorithms. *Nat. Hazards* **2022**, *110*, 899–935. [[CrossRef](#)]

42. Vilar del Hoyo, L.; Martín Isabel, M.P.; Martínez Vega, F.J. Logistic Regression Models for Human-Caused Wildfire Risk Estimation: Analysing the Effect of the Spatial Accuracy in Fire Occurrence Data. *Eur. J. For. Res.* **2011**, *130*, 983–996. [[CrossRef](#)]
43. Jafari Goldarag, Y.; Mohammadzadeh, A.; Ardakani, A.S. Fire Risk Assessment Using Neural Network and Logistic Regression. *J. Indian Soc. Remote Sens.* **2016**, *44*, 885–894. [[CrossRef](#)]
44. Mohammadi, F.; Bavaghar, M.P.; Shabaniyan, N. Forest Fire Risk Zone Modeling Using Logistic Regression and GIS: An Iranian Case Study. *Small-Scale For.* **2014**, *13*, 117–125. [[CrossRef](#)]
45. Milanović, S.; Marković, N.; Pamučar, D.; Gigović, L.; Kostić, P.; Milanović, S.D. Forest Fire Probability Mapping in Eastern Serbia: Logistic Regression versus Random Forest Method. *Forests* **2021**, *12*, 5. [[CrossRef](#)]
46. Sivrikaya, F.; Küçük, Ö. Modeling Forest Fire Risk Based on GIS-Based Analytical Hierarchy Process and Statistical Analysis in Mediterranean Region. *Ecol. Inform.* **2022**, *68*, 101537. [[CrossRef](#)]
47. Pourghasemi, H.; Beheshtirad, M.; Pradhan, B. 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. Hazards Risk* **2016**, *7*, 861–885. [[CrossRef](#)]
48. Kayet, N.; Chakrabarty, A.; Pathak, K.; Sahoo, S.; Dutta, T.; Hatai, B.K. Comparative Analysis of Multi-Criteria Probabilistic FR and AHP Models for Forest Fire Risk (FFR) Mapping in Melghat Tiger Reserve (MTR) Forest. *J. For. Res.* **2020**, *31*, 565–579. [[CrossRef](#)]
49. Mohajane, M.; Costache, R.; Karimi, F.; Bao Pham, Q.; Essahlaoui, A.; Nguyen, H.; Laneve, G.; Oudija, F. Application of Remote Sensing and Machine Learning Algorithms for Forest Fire Mapping in a Mediterranean Area. *Ecol. Indic.* **2021**, *129*, 107869. [[CrossRef](#)]
50. Tariq, A.; Shu, H.; Siddiqui, S.; Munir, I.; Sharifi, A.; Li, Q.; Lu, L. Spatio-Temporal Analysis of Forest Fire Events in the Margalla Hills, Islamabad, Pakistan Using Socio-Economic and Environmental Variable Data with Machine Learning Methods. *J. For. Res.* **2022**, *33*, 183–194. [[CrossRef](#)]
51. Shao, Y.; Feng, Z.; Sun, L.; Yang, X.; Li, Y.; Xu, B.; Chen, Y. Mapping China's Forest Fire Risks with Machine Learning. *Forests* **2022**, *13*, 856. [[CrossRef](#)]
52. Raduła, M.W.; Szymura, T.H.; Szymura, M.; Swacha, G. Macroecological Drivers of Vascular Plant Species Composition in Semi-natural Grasslands: A Regional Study from Lower Silesia (Poland). *Sci. Total Environ.* **2022**, *833*, 155151. [[CrossRef](#)]
53. Karger, D.N.; Conrad, O.; Böhrner, J.; Kawohl, T.; Kreft, H.; Soria-Auza, R.W.; Zimmermann, N.E.; Linder, H.P.; Kessler, M. Climatologies at High Resolution for the Earth's Land Surface Areas. *Sci. Data* **2017**, *4*, 170122. [[CrossRef](#)]
54. Milewski, W. *Forests in POLAND 2018*; The State Forests Information Centre: Warsaw, Poland, 2018; ISBN 978-83-65659-40-8.
55. Carmo, M.; Moreira, F.; Casimiro, P.; Vaz, P. Land Use and Topography Influences on Wildfire Occurrence in Northern Portugal. *Landsc. Urban Plan.* **2011**, *100*, 169–176. [[CrossRef](#)]
56. Konkathi, P.; Shetty, A.; Kolluru, V.; Yathish, P.H.; Pruthviraj, U. Static Fire Risk Index for the Forest Resources of Karnataka. In Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 6716–6719. [[CrossRef](#)]
57. Ye, T.; Wang, Y.; Guo, Z.; Li, Y. Factor Contribution to Fire Occurrence, Size, and Burn Probability in a Subtropical Coniferous Forest in East China. *PLoS ONE* **2017**, *12*, e0172110. [[CrossRef](#)]
58. Jaafari, A.; Gholami, D.M.; Zenner, E.K. A Bayesian Modeling of Wildfire Probability in the Zagros Mountains, Iran. *Ecol. Inform.* **2017**, *39*, 32–44. [[CrossRef](#)]
59. Guo, F.; Wang, G.; Su, Z.; Liang, H.; Wang, W.; Lin, F.; Liu, A. What Drives Forest Fire in Fujian, China? Evidence from Logistic Regression and Random Forests. *Int. J. Wildland Fire* **2016**, *25*, 505–519. [[CrossRef](#)]
60. Nhungo, E.J.S.; Fontana, D.C.; Guasselli, L.A.; Bremm, C. Probabilistic Modelling of Wildfire Occurrence Based on Logistic Regression, Niassa Reserve, Mozambique. *Geom. Nat. Hazards Risk* **2019**, *10*, 1772–1792. [[CrossRef](#)]
61. Ciesielski, M.; Bałazy, R.; Borkowski, B.; Szczyński, W.; Zasada, M.; Kaczmarowski, J.; Kwiatkowski, M.; Szczygieł, R.; Milanović, S. Contribution of Anthropogenic, Vegetation, and Topographic Features to Forest Fire Occurrence in Poland. *iForest* **2022**, *15*, 307–314. [[CrossRef](#)]
62. Kolecka, N.; Kozak, J. Assessment of the Accuracy of SRTM C- and X-Band High Mountain Elevation Data: A Case Study of the Polish Tatra Mountains. *Pure Appl. Geophys.* **2014**, *171*, 897–912. [[CrossRef](#)]
63. Oliver, M.A.; Webster, R. Kriging: A Method of Interpolation for Geographical Information Systems. *Int. J. Geogr. Inf. Syst.* **1990**, *4*, 313–332. [[CrossRef](#)]
64. Midi, H.; Sarkar, S.K.; Rana, S. Collinearity Diagnostics of Binary Logistic Regression Model. *J. Interdiscip. Math.* **2010**, *13*, 253–267. [[CrossRef](#)]
65. Kuhn, M.; Johnson, K. Data Preprocessing. In *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013; pp. 27–59. [[CrossRef](#)]
66. Hall, P.; Kurka, M.; Bartz, A. *Using H₂O Driverless Ai*; H2O.ai, Inc.: Mountain View, CA, USA, 2022; pp. 1–136.
67. Hosmer, D.W.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*, 3rd ed.; Wiley: Hoboken, NJ, USA, 2013; ISBN 9781118548387.
68. López-Ratón, M.; Rodríguez-Álvarez, M.X.; Suárez, C.C.; Sampedro, F.G. OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests. *J. Stat. Softw.* **2014**, *61*, 1–36. [[CrossRef](#)]

69. Goksuluk, D.; Korkmaz, S.; Zararsiz, G.; Karaagaoglu, A. easyROC: An Interactive Web-Tool for ROC Curve Analysis Using R Language Environment. *R J.* **2016**, *8*, 213–230. [[CrossRef](#)]
70. Romero, R.; Mestre, A.; Botey, R. A New Calibration for Fire Weather Index in Spain (AEMET). In *Advanced Forest Fire Research*; Imprensa da Universidade de Coimbra: Coimbra, Portugal, 2014; pp. 1044–1053. [[CrossRef](#)]
71. Austin, P.C.; Steyerberg, E.W. Events per Variable (EPV) and the Relative Performance of Different Strategies for Estimating the Out-of-Sample Validity of Logistic Regression Models. *Stat. Methods Med. Res.* **2017**, *26*, 796–808. [[CrossRef](#)]
72. Bujang, M.A.; Sa'at, N.; Sidik, T.M.I.T.A.B.; Joo, L.C. Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy between Statistics and Parameters Based on Real Life Clinical Data. *Malays. J. Med. Sci.* **2018**, *25*, 122–130. [[CrossRef](#)]
73. Varner, J.M.; Kane, J.M.; Kreye, J.K.; Engber, E. The Flammability of Forest and Woodland Litter: A Synthesis. *Curr. For. Rep.* **2015**, *1*, 91–99. [[CrossRef](#)]
74. Xanthopoulos, G.; Calfapietra, C.; Fernandes, P. Fire Hazard and Flammability of European Forest Types. In *Post-Fire Management and Restoration of Southern European Forests. Managing Forest Ecosystems*; Moreira, F., Arianoutsou, M., Corona, P., De las Heras, J., Eds.; Springer: Dordrecht, The Netherlands, 2012; Volume 24. [[CrossRef](#)]
75. Galiana-Martin, L.; Herrero, G.; Solana, J. A Wildland-Urban Interface Typology for Forest Fire Risk Management in Mediterranean Areas. *Landsc. Res.* **2011**, *36*, 151–171. [[CrossRef](#)]
76. Chappaz, F.; Ganteaume, A. Role of Land-Cover and WUI Types on Spatio-temporal Dynamics of Fires in the French Mediterranean Area. *Risk Anal.* **2022**, *2022*, 1–26. [[CrossRef](#)] [[PubMed](#)]
77. Fox, D.M.; Martin, N.; Carrega, P.; Andrieu, J.; Adnès, C.; Emsellem, K.; Ganga, O.; Moebius, F.; Tortorollo, N.; Fox, E.A. Increases in Fire Risk Due to Warmer Summer Temperatures and Wildland Urban Interface Changes Do Not Necessarily Lead to More Fires. *Appl. Geogr.* **2015**, *56*, 1–12. [[CrossRef](#)]
78. Molina, J.R.; Ortega, M.; Rodríguez, Y.; Silva, F.F. Fire Ignition Patterns to Manage Prescribed Fire Behavior: Application to Mediterranean Pine Forests. *J. Environ. Manag.* **2022**, *302*, 114052. [[CrossRef](#)]
79. Benali, A.; Sá, A.C.L.; Ervilha, A.R.; Trigo, R.M.; Fernandes, P.M.; Pereira, J.M.C. Fire Spread Predictions: Sweeping Uncertainty under the Rug. *Sci. Total Environ.* **2017**, *592*, 187–196. [[CrossRef](#)]
80. Atchley, A.L.; Linn, R.; Jonko, A.; Hoffman, C.; Hyman, J.D.; Pimont, F.; Sieg, C.; Middleton, R.S. Effects of Fuel Spatial Distribution on Wildland Fire Behaviour. *Int. J. Wildland Fire* **2021**, *30*, 179–189. [[CrossRef](#)]
81. Clements, C.B.; Lareau, N.P.; Seto, D.; Contezac, J.; Davis, B.; Teske, C.; Zajkowski, T.J.; Hudak, A.T.; Bright, B.C.; Dickinson, M.B.; et al. Fire Weather Conditions and Fire—Atmosphere Interactions Observed during Low-Intensity Prescribed Fires—RxCADRE 2012. *Int. J. Wildland Fire* **2016**, *25*, 90–101. [[CrossRef](#)]
82. Anderson, H.E. Forest Fuel Ignitability. *Fire Technol.* **1970**, *6*, 312–319. [[CrossRef](#)]
83. Hilton, J.E.; Miller, C.; Sullivan, A.L.; Rucinski, C. Effects of Spatial and Temporal Variation in Environmental Conditions on Simulation of Wildfire Spread. *Environ. Modell. Softw.* **2015**, *67*, 118–127. [[CrossRef](#)]
84. Balzter, H.; Gerard, F.; George, C.; Weedon, G.; Grey, W.; Combal, B.; Bartholomé, E.; Bartalev, S.; Los, S. Coupling of Vegetation Growing Season Anomalies and Fire Activity with Hemispheric and Regional-Scale Climate Patterns in Central and East Siberia. *J. Clim.* **2007**, *20*, 3713–3729. [[CrossRef](#)]
85. Amiro, B.D.; Stocks, B.J.; Alexander, M.E.; Flannigan, M.D.; Wotton, B.M. Fire, Climate Change, Carbon and Fuel Management in the Canadian Boreal Forest. *Int. J. Wildland Fire* **2001**, *10*, 405–413. [[CrossRef](#)]
86. Satir, O.; Berberoglu, S.; Donmez, C. Mapping Regional Forest Fire Probability Using Artificial Neural Network Model in a Mediterranean Forest Ecosystem. *Geomat. Nat. Hazards Risk* **2016**, *7*, 1645–1658. [[CrossRef](#)]
87. Leuenberger, M.; Parente, J.; Tonini, M.; Pereira, M.G.; Kanevski, M. Wildfire Susceptibility Mapping: Deterministic vs. Stochastic Approaches. *Environ. Modell. Softw.* **2018**, *101*, 194–203. [[CrossRef](#)]
88. Cao, Y.; Wang, M.; Liu, K. Wildfire Susceptibility Assessment in Southern China: A Comparison of Multiple Methods. *Int. J. Disaster Risk Sci.* **2017**, *8*, 164–181. [[CrossRef](#)]
89. Trucchia, A.; Meschi, G.; Fiorucci, P.; Gollini, A.; Negro, D. Defining Wildfire Susceptibility Maps in Italy for Understanding Seasonal Wildfire Regimes at the National Level. *Fire* **2022**, *5*, 30. [[CrossRef](#)]
90. Nguyen, Q.-H.; Nguyen, H.-D.; Le, D.T.; Bui, Q.-T. Fine-Tuning LightGBM Using an Artificial Ecosystem-Based Optimizer for Forest Fire Analysis. *Forest Sci.* **2022**, *2022*, fxac039. [[CrossRef](#)]
91. Mansoor, S.; Farooq, I.; Kachroo, M.M.; Mahmoud, A.E.D.; Fawzy, M.; Popescu, S.M.; Alyemeni, M.N.; Sonne, C.; Rinklebe, J.; Ahmad, P. Elevation in Wildfire Frequencies with Respect to the Climate Change. *J. Environ. Manag.* **2022**, *301*, 113769. [[CrossRef](#)]
92. Ruiz González, A.D.; Vega Hidalgo, J.A.; Álvarez González, J.G. Construction of Empirical Models for Predicting Pinus sp. Dead Fine Fuel Moisture in NW Spain. I: Response to Changes in Temperature and Relative Humidity. *Int. J. Wildland Fire* **2009**, *18*, 71–83. [[CrossRef](#)]
93. Matthews, S. A Process-Based Model of Fine Fuel Moisture. *Int. J. Wildland Fire* **2006**, *15*, 155–168. [[CrossRef](#)]
94. Aragão, L.E.O.C.; Malhi, Y.; Roman-Cuesta, R.M.; Saatchi, S.; Anderson, L.O.; Shimabukuro, Y.E. Spatial Patterns and Fire Response of Recent Amazonian Droughts. *Geophys. Res. Lett.* **2007**, *34*, L07701. [[CrossRef](#)]
95. Chuvieco, E.; Aguado, I.; Dimitrakopoulos, A.P. Conversion of Fuel Moisture Content Values to Ignition Potential for Integrated Fire Danger Assessment. *Can. J. For. Res.* **2004**, *34*, 2284–2293. [[CrossRef](#)]

96. Chang, Y.; He, H.S.; Hu, Y.; Bu, R.; Li, X. Historic and Current Fire Regimes in the Great Xing'an Mountains, Northeastern China: Implications for Long-Term Forest Management. *For. Ecol. Manag.* **2008**, *254*, 445–453. [[CrossRef](#)]
97. Khabarov, N.; Krasovskii, A.; Obersteiner, M.; Swart, R.; Dosio, A.; San-Miguel-Ayanz, J.; Durrant, T.; Camia, A.; Migliavacca, M. Forest Fires and Adaptation Options in Europe. *Reg. Environ. Chang.* **2016**, *16*, 21–30. [[CrossRef](#)]
98. Agee, J.K.; Skinner, C.N. Basic Principles of Forest Fuel Reduction Treatments. *For. Ecol. Manag.* **2005**, *211*, 83–96. [[CrossRef](#)]
99. Agee, J.K.; Bahro, B.; Finney, M.A.; Omi, P.N.; Sapsis, D.B.; Skinner, C.N.; van Wagendonk, J.W.; Phillip Weatherspoon, C. The Use of Shaded Fuelbreaks in Landscape Fire Management. *For. Ecol. Manag.* **2000**, *127*, 55–66. [[CrossRef](#)]
100. Kolanek, A.; Szymanowski, M.; Raczyk, A. Human Activity Affects Forest Fires: The Impact of Anthropogenic Factors on the Density of Forest Fires in Poland. *Forests* **2021**, *12*, 728. [[CrossRef](#)]
101. Hysa, A.; Spalevic, V.; Dudic, B.; Roşca, S.; Kuriqi, A.; Bilaşco, Ş.; Sestras, P. Utilizing the Available Open-Source Remotely Sensed Data in Assessing the Wildfire Ignition and Spread Capacities of Vegetated Surfaces in Romania. *Remote Sens.* **2021**, *13*, 2737. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.