

THÈSE DE DOCTORAT

Évaluation de la dynamique spatiotemporelle des incendies de forêt dans le sud de la France

Annexes

Christos BOUNTZOUKLIS

ESPACE - Etude des Structures, des Processus d'Adaptation et des
Changements de l'Espace

**Présentée en vue de l'obtention
du grade de docteur en Géographie
d'Université Côte d'Azur**
Dirigée par : Dennis Fox
Co-encadrée par : Elena Di Bernardino
Soutenue le : 5 juillet 2023

Devant le jury, composé de :
Etienne Cossart, Professeur des
universités, Université Jean Moulin Lyon 3
Lea Wittenberg, Professeure, University of
Haifa
Marielle Jappiot, Ingénieure de Recherche,
INRAE
Dennis Fox, Professeur des universités,
Université Côte d'Azur
Elena Di Bernardino, Professeure des
universités, Université Côte d'Azur



Environmental factors affecting wildfire-burned areas in southeastern France, 1970–2019

Christos Bountzouklis¹, Dennis M. Fox¹, and Elena Di Bernardino²

¹UMR CNRS 7300 ESPACE, Université Côte d'Azur, Nice 06204, France

²UMR CNRS 7351 LJAD, Université Côte d'Azur, Nice 06108, France

Correspondence: Christos Bountzouklis (christos.bountzouklis@univ-cotedazur.fr)

Received: 16 June 2021 – Discussion started: 28 June 2021

Revised: 1 March 2022 – Accepted: 7 March 2022 – Published: 6 April 2022

Abstract. Forest fires burn an average of about 440 000 ha each year in southern Europe. These fires cause numerous casualties and deaths and destroy houses and other infrastructure. In order to elaborate on suitable firefighting strategies, complex interactions between human and environmental factors must be taken into account. In this study, we investigated the spatiotemporal evolution in the burned area over a 50-year period (1970–2019) and its interactions with topography (slope aspect and inclination) and vegetation type in southeastern France by exploiting the geographic information system (GIS) databases. Data were analyzed for two 25-year periods (1970–1994 and 1995–2019), since a new fire suppression policy was put into place after 1994, which focused on rapid extinction of fires in their early phase. In the last 25 years, the burned area decreased sharply, and the geographic distribution of fires also changed, especially in regions where large fires occur (Var administrative division). Elsewhere, even though forest fires remain frequent, the total extent of the burned area decreased substantially. Fire hotspots appear closer to built-up areas in the west, are randomly distributed in the east, and they almost completely disappear in the central region of the study area where there is a history of large fires. Slope orientation presents an increasingly important role in the second period; south-facing slopes are preferred the most by fire, and north-facing slopes are preferentially avoided. Even though the slope inclination is less affected by the new firefighting strategy, low slope inclinations are even more avoided after 1994. The greatest proportion of the burned area is strongly associated with the location of sclerophyllous vegetation clusters which exhibit highly fire prone and expand in area over time. Natural grass-

lands are also preferred by fire, while broadleaved, coniferous, and mixed forest are increasingly avoided by fire.

1 Introduction

Forest fire is a common and important element of the Earth system (Bond and Keeley, 2005) that disturbs natural ecosystems and threatens human welfare and wellbeing throughout the globe. The Mediterranean climate is characterized by hot and dry summers which favor fire ignition and propagation. Consequently, wildfires are particularly active around the Mediterranean basin, and fires in the Mediterranean climate zones are considered to have a wide range of environmental and socioeconomic impacts (Miller et al., 2009; San-Miguel-Ayanz et al., 2013; Ganteaume et al., 2013).

Forest fires burn an average of 440 000 ha each year in the Euro–Mediterranean region, and this corresponds to about 85 % of the total burned area (BA) in Europe (San-Miguel-Ayanz et al., 2020). Of the five principal Euro–Mediterranean countries concerned by forest fires (France, Greece, Italy, Portugal, and Spain), France has the lowest amount of BA (San-Miguel-Ayanz et al., 2020). It also has the smallest potential burnable area, since only the southern Mediterranean fringe is affected by forest fires. France, Spain, Italy, and Greece all show similar trends in decreasing decadal BA in 1980–2010, and only Portugal experienced a progressive increase during this interval (San-Miguel-Ayanz et al., 2020). It should be noted that BA is generally decreasing despite increases in summer temperatures throughout the Euro–Mediterranean zone (Pokorná et al., 2018; Rodrigues et al., 2020), and this can be attributed to more efficient firefighting

strategies (Fox et al., 2015; Turco et al., 2016; Ganteaume and Barbero, 2019).

Forest fire spatial distribution, size, and frequency are associated with several interacting factors that can be categorized into two main groups, namely (i) environmental and (ii) anthropogenic factors. Environmental factors generally include fuel characteristics (e.g., type and water content), topography (e.g., slope inclination, altitude, and aspect), and weather conditions (e.g., temperature, humidity, and wind speed); anthropogenic factors include the characteristics of the transitional zone between wildland vegetation and artificial areas in the wildland–urban interface (WUI).

Among the environmental characteristics, several studies provide evidence of spatial patterns relating forest fire probability and BA to topography (Dickson et al., 2006; Nunes et al., 2016; Padilla and Vega-García, 2011). Slope aspect affects incoming solar radiation and can determine fuel type, fuel moisture, and fuel density, all of which influence flammability (Holden et al., 2009). In addition, aspect influences the degree of ecological change related to fire (fire severity; Birch et al., 2015; Estes et al., 2017; Parks et al., 2018). In the Northern Hemisphere, south-facing slopes receive more solar radiation during the day than north-facing slopes, and this can enhance burn severity (Alexander et al., 2006; Oliveira et al., 2014a; Oliveras et al., 2009), but the trend is not systematic (Broncano and Retana, 2004). In addition to the impact on fire severity, other studies (Mouillot et al., 2003) have demonstrated that south-facing slopes in Corsica (France) can burn more frequently than other exposures. On the northern shore of the Mediterranean, south-facing slopes frequently have more housing than north-facing slopes, and this may contribute to a greater number of ignitions (Fox et al., 2018). Steep slopes tend to have higher spread rates and fire intensities (Capra et al., 2018); fatality rates are also greater compared to flat areas (Molina-Terrén et al., 2019). Csontos and Cseresnyés (2015) observed an exponential increase in upslope fire spread with increase in slope inclination, whereas downslope fire spread velocity was unaffected by slope angle and was similar to rates detected on flat terrain. Slope and altitude tend to be correlated, but their association with fires is often conflicting. For instance, Nunes et al. (2016) found that BA and ignition density were positively correlated with elevation and slope at a municipal scale in Portugal. Similarly, Elia et al. (2019) showed that the probability of fire ignition increased with elevation and slope in southern Italy. However, Narayanaraj and Wimberly (2012) observed a negative impact of elevation and slope inclination on human-caused fires.

The role of vegetation is complex and can be influenced by flammability (Michelaki et al., 2020; Molina et al., 2017) or spatial patterns of vegetation in the landscape (Curt et al., 2013). Vegetation continuity affects the fire propagation, which contributes to determining the BA (Duane et al., 2015; Fernandes et al., 2016). Vegetation type is another important factor to consider which has been explored in number of

studies on fire selectivity indices (Bajocco and Ricotta, 2008; Barros and Pereira, 2014; Carmo et al., 2011; Moreira et al., 2009; Moreno et al., 2011; Nunes et al., 2005; Pereira et al., 2014). Overall, there is widespread agreement in the literature that shrublands are regarded as fire prone areas at multiple scales, including regional (Carmo et al., 2011; Moreno et al., 2011), national (Nunes et al., 2016, 2005), and continental (Moreira et al., 2011; Oliveira et al., 2014b; Pereira et al., 2014) scales. The probability of large fires is greater in dense shrublands than in forested ecosystems in the Mediterranean basin (Moreira et al., 2011; Ruffault and Mouillot, 2017). According to Mermoz et al. (2005), fire proneness of shrublands could be related to their recovery rate, since shrublands can regenerate faster and favor fuel accumulation in a short time, unlike forests which take longer to recover and expand. In addition, Oehler et al. (2012) point out that shrubs are considered a low suppressing priority by firefighters due to the low cost of restoration. In Europe, grasslands are also considered to be fire prone (Oliveira et al., 2014a). Cultivated areas are the least fire prone vegetation types because of their low combustibility and proximity to built-up land covers, which facilitates rapid fire detection and suppression (Moreira et al., 2011). Forested areas are found to be more fire prone than cultivated areas but less than shrublands (Moreira et al., 2011). More specifically, broadleaved forests are usually less prone to burning than coniferous species, which present a greater fire hazard (Moreira et al., 2009; Oliveira et al., 2014a).

Spatial relationships between fire occurrence and environmental factors evolve over time, due to changes in biomes and climate, but also as the result of fire management practices. Mapping and understanding these trends are crucial for evaluating the effectiveness of firefighting strategies and developing suitable policies (Bowman et al., 2017). There are numerous recent efforts that aim to analyze the spatial and temporal trends of fire activity at a global, national, and regional level. Otón et al. (2021) analyzed global trends of BA, based on the FireCCILT11 database, which is the longest available global BA dataset to date (1982–2018). At a national level, Catarino et al. (2020) investigated the trends of annual BA in Angola between 2001 and 2019 using MODIS products (MCD64A1) and associated the significant trends to land cover, ecological regions, and protected areas. Ganteaume and Barbero (2019) utilized a long-term (1957–2017) fire geodatabase to analyze the spatiotemporal variations in large fires in terms of frequency and BA in the French Mediterranean. Silva et al. (2019) used a satellite-derived BA dataset covering a 39-year period over the Iberian Peninsula to study BA trends and explore the relationship between areas with significant BA trends and fire danger. Urbieta et al. (2019) studied the spatiotemporal trends in Spain between 1980 and 2013 with regard to fire frequency, BA, and fire size and their relationship with changes in climate, land use and land cover, and fire suppression. Viedma et al. (2018) assessed the changing role of environmental and human-related

factors in reference to fire activity in west-central Spain from 1979 to 2008. Fire suppression is an important factor that can influence fire spread. In France, as a response to the large fires that occurred between 1986 and 1990, a major change in the fire suppression strategy was established in the 1990s; it focused on the rapid suppression of fire ignitions, regardless of the weather conditions, in order to avoid fire propagation (Battesti, 1997). The fire policy had a significant impact on fire activity in southern France and weakened the fire–weather relationship (Ruffault and Mouillot, 2015). Despite the sharp decrease in BA after the full implementation of the fire management policy, its effectiveness on very large fires was not as successful as for smaller fires, since changes in BA that correspond to large return periods are not significant (Evin et al., 2018). Although many studies have focused on determining the relationships between fire behavior and driving factors (Mhaweji et al., 2015), few studies have examined how fire suppression strategies impact the spatial distribution of BA. Identifying spatial patterns and the main driving forces that determine fire distribution provides useful information for fire and civil protection agencies, and it assists in the allocation of appropriate firefighting resources and in designing proper prevention actions (Moreira et al., 2011).

The objective of this study is to quantify changes in spatiotemporal BA patterns induced by a major shift in the fire suppression strategy initiated in the early 1990s in southeastern France. The time interval under study spans 5 decades (1970–2019) and includes the relation of BA with respect to environmental factors such as (a) topography (slope aspect and inclination) and (b) vegetation type. Although several studies have investigated the relationships between BA and environmental factors, very few have covered such a long time interval based on burn scar polygons, and they have not been explicitly related to changes in fire suppression methods.

2 Data and methods

2.1 Study area

The study area is comprised of a subset of the three administrative division, with the greatest BA in continental France (only Corsica has a greater burned area), according to the French official forest fire database (<http://promethee.com>, last access: 15 February 2022) for Bouches-du-Rhône, Var, and Alpes-Maritimes (Table 1; Fig. 1). Areas within the administrative division limits that were excluded represent surfaces that cannot burn, such as marshlands in the westernmost part of Bouches-du-Rhône and the high alpine mineral surfaces located in the northern part of Alpes-Maritimes.

Topography varies from west to east (Fig. 1). The gentlest slopes are found in the west (Bouches-du-Rhône), and both the altitude and slope inclination increase eastwards. The steepest slope inclinations are found in the northeastern

Table 1. Environmental characteristics of the study area per administrative division unit.

	Bouches-du-Rhône	Var	Alpes-Maritimes
Total area (km ²)	3456	6019	3495
Forested area (km ²)	1530	4044	2727
Ratio forest/total (km ²)	0.44	0.67	0.78
Mean slope (°)	8.8	11.9	24.3
Median slope (°)	5.7	9.6	25.2

part of the study area where the French Alps are located. Topography influences the population distribution since much of the built area is concentrated along the coast or on shallow to intermediate slopes in the WUI. In the Bouches-du-Rhône area, the western portion of the administrative division has particularly low population densities due to the presence of the national park and wetlands mentioned above. Similarly, much of the population in the Alpes-Maritimes region is concentrated in the southern portion of the administrative division. The 2010 population densities of 388.8, 167.5, and 252.0 persons per square kilometer for the Bouches-du-Rhône, Var, and Alpes-Maritimes, respectively, are approximate as they simply divide population by total area without accounting for geographic distributions. The order, however, is accurate and shows the greatest population density for Bouche-du-Rhône, and the lowest for the Var. Based on the demographic and environmental characteristics described above, the westernmost section (Bouche-du-Rhône) of the study area has low potential for fire ignition and propagation, but this increases when moving towards the eastern half of administrative division. The central part of the study area (Var administrative division) has a high potential for fire ignition and the greatest potential for fire propagation, since it has a high forested area and a large continuous WUI area. Finally, the eastern section (Alpes-Maritimes administrative division) has high ignition and propagation potentials in the southern portion of the administrative division and low ignition/high propagation at higher altitudes.

2.2 Fire database

Forest fire research in France is frequently based on the national database for forest fires in France (<http://promethee.com>), where the fire location is defined as the municipality where fire ignition occurred. For this study, we used a fire geographic information systems (GIS) database provided by the National Forestry Office (Office National des Forêts – ONF) and the Delegation for the Protection of the Mediterranean Forest (Délégation à la Protection de la Forêt Méditerranéenne – DPFM). Even though the number of recorded fires is significantly lower than the Prométhée database, the total area burned is almost identical; the very small fires recorded in Prométhée are not all digitized in the



Figure 1. Map of southeastern France showing the study area and the administrative division limits overlain on a 5 m digital elevation model.

ONF database. To the best of our knowledge, this is only the second use of this geodatabase after Ganteaume and Barbero (2019). The dataset includes more than 3000 digitized burn scar polygons for fires that occurred between 1970 and 2019. Due to the long temporal extent of the database, the accuracy and the methods used to define burn scars varied over time. In the 1970s, burn scars were mapped using field measurements with GPS devices, and the technique progressively evolved to integrate remote sensing data (satellite imagery and orthophotos). Although the description of how BA was defined is not recorded in the database, earlier polygons are clearly less accurate (coarse shapes with little detail) than burn scars after the advent of satellite imagery (Fig. 2).

2.3 Environmental variables

2.3.1 Topography

Burn scar polygons were rasterized to a 5 m spatial resolution and overlain on a 5 m digital elevation model (DEM) extracted from RGE ALTI©, the official National Geographic Institute (Institut Géographique National – IGN) database. The DEM was used to calculate the slope aspect and inclination. In the conversion of vector polygons to raster cells, BA polygons smaller than half the cell size (25 m^2) were not defined as being burned during rasterization, so BA for the slope aspect and inclination analyses represent approximately 96 % of actual BA in the study area. The aspect was divided into the following five categories: flat, north, east, south, and west. The inclination was divided into the following five categories: 0–10, 10–20, 20–30, 30–40, and $> 40^\circ$.

Table 2. CORINE Land Cover layers and their respective fire periods.

CORINE Land Cover	Fire period
1972 (predicted)	1970–1974
1980 (predicted)	1975–1984
1990	1985–1994
2000	1995–2002
2006	2003–2009
2012	2010–2014
2018	2015–2019

2.3.2 Vegetation type

For the computation of the forested BA and the identification of fire prone vegetation categories, GIS forest layers were extracted from the European CORINE Land Cover (CLC) database. The database includes five reference years of 1990, 2000, 2006, 2012, and 2018. In addition to the CLC reference layers, it was considered best to backcast two additional forest cover layers for 1972 and 1980 to account for any transitions between forested and non-forested surfaces for the 2 decades preceding the CLC database. The methodology followed for the projection process is addressed in the section on forest layer projection. The fire geodatabase was then matched with the CLC layer that was chronologically closest to the equivalent fire period (see Table 2).

The vegetation types that were used in the current study follow the CLC nomenclature, i.e., broadleaved forest, coniferous forest, mixed forest, natural grasslands, and sclerophyllous vegetation (Fig. 3; Table 3). Although natural grasslands and sclerophyllous vegetation are not forests, the categories

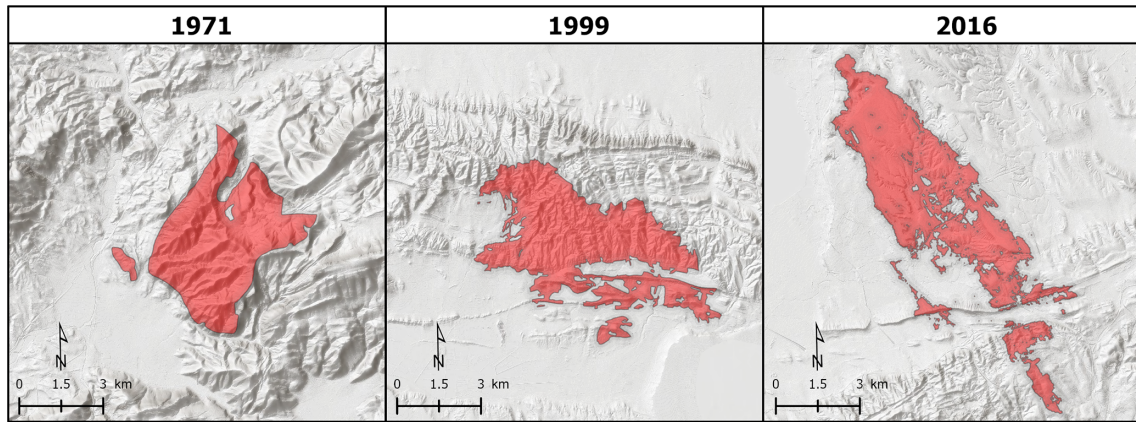


Figure 2. Evolution of digitized burn scar accuracy over the past few decades.

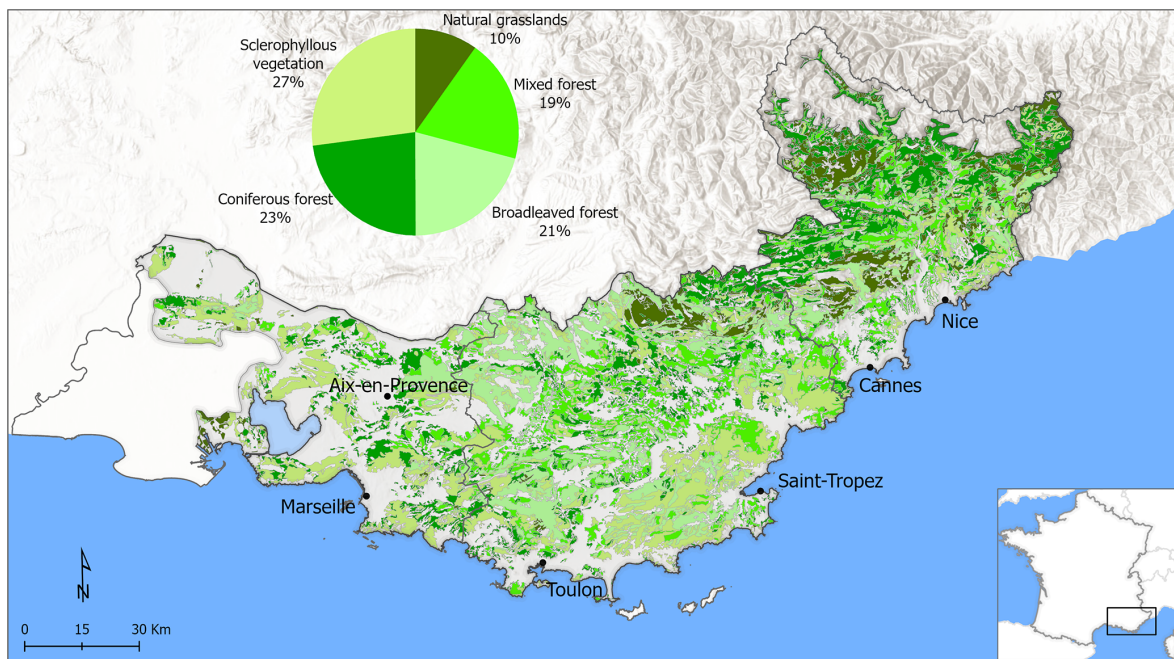


Figure 3. Distribution of vegetation types based on CLC 2018.

Table 3. Average and relative forested areas according to vegetation type between 1970 to 2019.

Type	Area (ha)	%
Broadleaved forest	172 547	20.6
Coniferous forest	201 262	24.1
Mixed forest	160 973	19.2
Natural grassland	93 322	11.2
Sclerophyllous vegetation	208 057	24.9
Total	836 161	

Forest layer projection

Although most urban growth occurred on agricultural land (Roy et al., 2015) and forest cover changed little, the Land Change Modeler (LCM) module of TerrSet (Eastman, 2020) was used to predict vegetation cover in 1972 and 1980. The LCM is programmed to forecast change from an earlier to a later date, so going back in time (a backcast) required the temporal inversion of the filenames for the 1990 (renamed to 2000) and 2000 (renamed to 1990) CLC layers; in this way, the land cover was simulated for 1980 and 1972. Land cover categories were simplified from the original CLC categories to the following: built, broadleaved forest (broad), coniferous forest (conifer), mixed forest, natural grasslands (grass),

will be referred to collectively as wildland or forested areas indiscriminately for the sake of brevity.

sclerophyllous vegetation (bush), other, and water. Only transitions greater than 0.05 % of the landscape (14.3 km²) were modeled, and these included the following: bush → grass, bush → other, built → other, grass → other, broad → bush, other → grass, bush → conifer, other → bush, bush → broad, bush → mixed, mixed → bush, other → conifer, mixed → broad, mixed → other, other → broad, other → mixed, broad → other, grass → bush, mixed → conifer, built → mixed, built → bush, and conifer → mixed. Note that these are the inverse of historical trends, so the built–mixed transition actually backcasts the historical transition of mixed forest to built area. The explanatory variables used to predict land cover change were the following: altitude, slope inclination, distance from built area, distance from broad, distance from conifer, distance from mixed, distance from grass, distance from bush, distance from other, and distance from water. According to Eastman (2020), Cramer's *V* values of ≥ 0.15 for explanatory variables are useful and should be kept in the model, and all explanatory variables used here met this criterion. Accuracy rates to model transitions ranged from 65 % to 90 %, with mean and median values of 78 % and 80 %, respectively.

2.4 Fire history 1970–2019

A 500 × 500 m grid (25 ha) was created and overlain on the study area in order to measure the percentage of each cell that was burned each year between 1970 and 2019 (50 years) (Fig. 4). These percentage values were then summed to produce the cumulative percentage of BA for each cell. This approach facilitated the effort to identify clusters of cells/areas that have been burned multiple times and to give an overview of the spatial distribution of BA in the region. To better illustrate the impact of suppression strategies on fire occurrence, the method was applied to two 25-year subsets of the fire dataset (i) 1970–1994, and (ii) 1995–2019 as the mid-point break corresponds to the major shift in firefighting strategy and allocated resources in France.

2.5 Spatiotemporal analysis – contextual Mann–Kendall

In order to identify spatiotemporal trends within the entire time period (1970–2019), a modified version of the Mann–Kendall test was applied (Kendall, 1975; Mann, 1945). The Mann–Kendall test is a non-parametric test which is used to statistically assess monotonic upward or downward trends for a variable through time. In this study, we used the contextual Mann–Kendall (CMK) test which was introduced by Neeti and Eastman (2011), and it differs from the original test by evaluating trends at a 3 × 3 cell neighborhood for each cell in a grid. The specific method has been used to assess trends in BA with satisfactory outcomes (Silva et al., 2019; Catarino et al., 2020; Otón et al., 2021). The CMK method was de-

vised from Tobler's first law of geography, which states that “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970). By assuming that trends show signs of spatial autocorrelation between adjacent cells, the CMK test allows for greater confidence in identifying the presence of a trend (Neeti and Eastman, 2011). However, it requires observations to be a set of independent random variables and, thus, applying the test on data that are temporally autocorrelated may lead to a false rejection of the null hypothesis of no trend (Douglas et al., 2000). To assess the temporal autocorrelation in our dataset, we applied the Durbin–Watson test (Durbin and Watson, 1950), and to remove it, the prewhitening procedure by Wang and Swail (2001), which preserves the same temporal trend but without the autocorrelation, was used (Fig. 5).

2.6 Fire selectivity (Jacobs' index)

In order to examine the fire proneness of the environmental variables considered in this study (slope aspect and inclination and vegetation type), a resource selection index was calculated for each 25-year interval (Fig. 7). Resource selection is based primarily on wildlife ecology (Manly et al., 2002), but its use has been extended to include fire selectivity (Bajocco and Ricotta, 2008; Barros and Pereira, 2014; Moreira et al., 2001, 2009; Moreno et al., 2011; Nunes et al., 2005; Oliveira et al., 2014a). The rationale behind fire selectivity is that fires burn selectively when the proportion of a class (e.g., type of vegetation) within a burned area is higher than the proportion of the available area to burn. The opposite applies when a specific class of variable is burned proportionally less than the available area (fire avoidance).

In our work, we used Jacobs' selectivity index (Jacobs, 1974) which is defined as follows:

$$D_i = \frac{r - p}{r + p - 2rp}. \quad (1)$$

r stands for the proportion of resource class *i* used by fire, and *p* is the proportion of resource class *i* available to fire. Jacobs' index values range between −1 and 1. Positive values indicate fire preference, and negative values indicate fire avoidance. The index was calculated for each class of the environmental factors (described in the subsequent sections) for each year. Similar to other studies (Barros and Pereira, 2014; Nunes et al., 2005), the available area for each fire to burn is defined as twice the amount of area burned by each fire (Fig. 6).

2.7 Geographically weighted regression

A geographically weighted regression (GWR) was used to quantify the impact of the change in the firefighting strategy on the relative importance of the environmental factors. GWR is applied in wide range of interdisciplinary fields, including forest fires (Koutsias et al., 2010; Martínez-Fernández et al., 2013; Nunes et al., 2016; Rodrigues et al.,

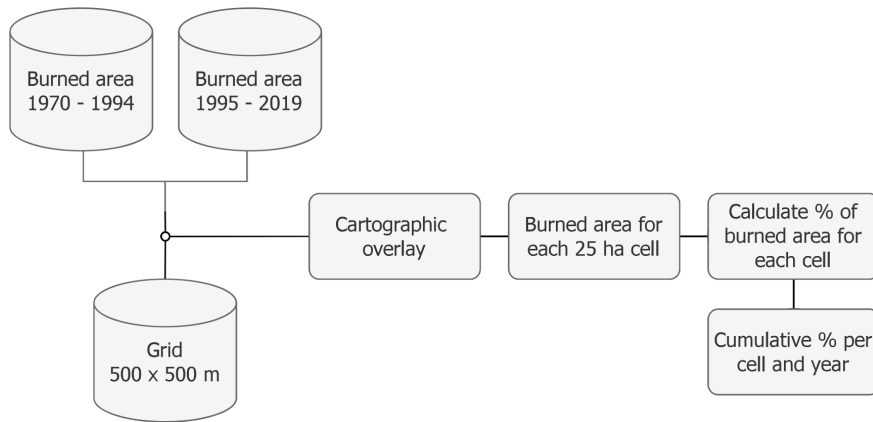


Figure 4. Flowchart depicting the processing steps for generating the cumulative percentage of forested burned area per cell.

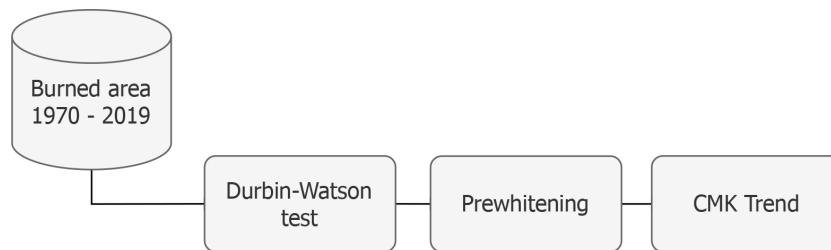


Figure 5. Flowchart depicting the processing steps for estimating the trend significance using the contextual Mann–Kendall method.

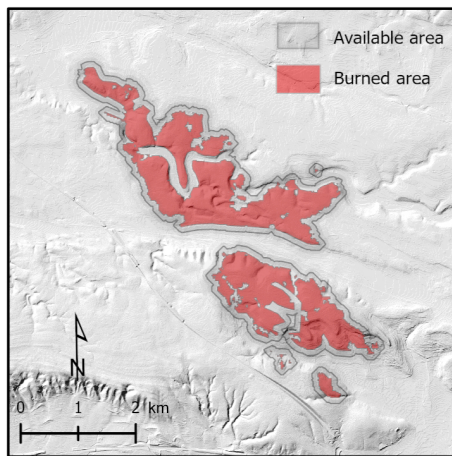


Figure 6. Illustration of the burned area (r) and available area (p) to be used by a fire. The available area (the sum of the burned area plus buffer zone) around each fire corresponds to twice the burned area.

2016; Kolanek et al., 2021). GWR is a local non-parametric regression method (Fotheringham et al., 2003) that allows the relationships between dependent and explanatory variables to vary over space. The basic form of a GWR model, provided by Fotheringham et al. (1998, 2003) is defined as follows:

$$y_i = \beta_{i0} + \sum_{z=1}^j \beta_{iz}x_{iz} + \varepsilon_i, \quad (2)$$

where y_i is the dependent variable at location i , β_{i0} is the intercept parameter at location i , j is the number of explanatory variables, β_{iz} is the local regression coefficient for the z th explanatory variable at location i , x_{iz} represents the z th explanatory variable at location i , and ε_i denotes the random error at location i . Since GWR allows coefficients to be spatially heterogeneous, a sub-model for the location of each observation is created that considers only a subsample of the total observations, where observations in closer proximity have a greater effect in determining the local set of coefficients than observations located at further distances (Fotheringham et al., 1998). This neighborhood is called a kernel, and the maximum distance from a regression point at a location i is defined as the bandwidth. The bandwidth is an important parameter than can be defined in the following two different ways: (i) fixed bandwidth (fixed distance for each regression point) and (ii) adaptive bandwidth (fixed number of nearest neighbors for each regression point). The first type of neighborhood is more appropriate when data are regularly distributed across space, whereas the second type is more appropriate for data that form spatial clusters. In the current work, the adaptive bandwidth approach was utilized to fit the GWR model which was optimized based on the value of Akaike information criterion (Akaike, 1974). For

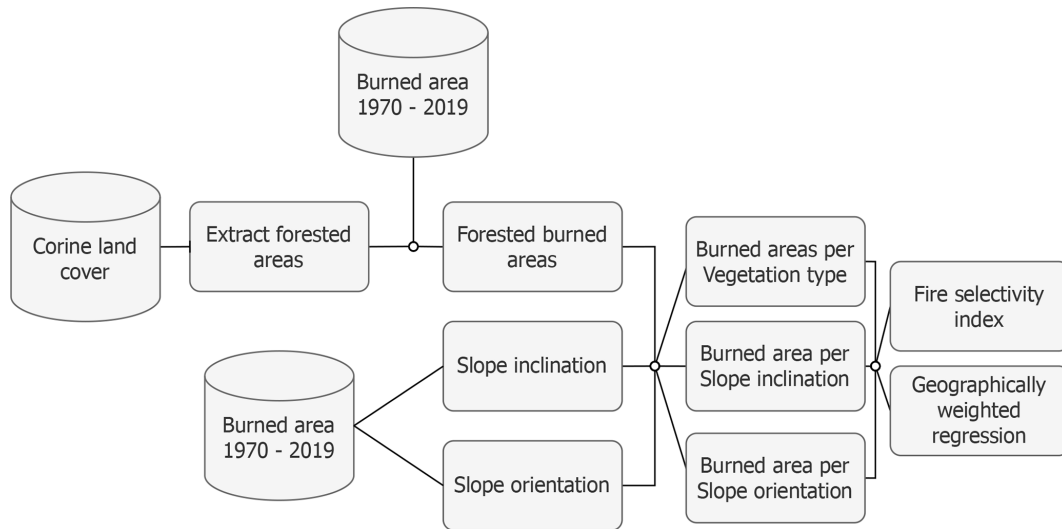


Figure 7. Flowchart depicting the processing steps and data used to relate the burned area to the vegetation type, slope inclination, and orientation.

each of the three environmental variables described above, a univariate GWR model was used to explore the relationship with the dependent variable (percent of BA) for the two 25-year periods of (i) 1970–1994 and (ii) 1995–2019.

3 Results

Results presented below will first describe the fire history for the entire time interval (1970–2019) and then analyze the spatiotemporal evolution of the BA split according to the two 25-year periods. Finally, it will explore the relationship of BA to topography (slope aspect and inclination) and vegetation type. Factor-specific results will be discussed, as they are presented in the following sections, while broader considerations will be explained in the Sect. 4.

3.1 Fire history

In total, 3382 fires burned 296 820 ha in 1970–2019. The mean and median areas of BA are 87.7 and 4.2 ha, respectively; these values reflect the typical positively skewed distribution of fire size, where the vast majority of fires are small, and a few fires, accounting for most of the burned area, are very large. The number of fires equal to or greater than 100, 500, and 1000 ha is 378 (11.2 %), 123 (3.6 %), and 65 (1.9 %), respectively. Of the total number of fires, 2424 (88.2 %) occurred in forested landscapes, and these burned an area of 263 645 ha (88.8 % of total BA).

Mean and median values for forested landscape fires are slightly greater than for all fires at 111.7 and 6.5 ha, respectively. The number of fires equal to or greater than 100, 500, and 1,000 ha is 314 (13.0 %), 106 (4.4 %), and 60 (2.5 %), respectively. As stated above, the results presented below will deal exclusively with the forested BA that was occupied by

one of the vegetation types mentioned in Sect. 2.3.2, since the trends with respect to vegetation and topography for all fires and forested landscapes are nearly identical.

Annual forested BA varies significantly from year to year (Fig. 8), although there are clear differences between the first 2 decades (1970–1990) and the last 3 decades (1991–2019). The mean and median annual BA are 5156.4 and 2746.1 ha, respectively. Several big fires occurred in the 1980s, followed by a sharp decrease in the early 1990s. Similar to the rest of southern Mediterranean Europe, most of the forested BA is related to a small number of large fires (Turco et al., 2016). Only 5 years (1979, 1986, 1989, 1990, and 2003) of the 50-year record account for almost half of the total forested BA (126 700 ha). The forested BA for each of these years surpasses 20 000 ha, attaining nearly 36 000 ha in 1989. Of the 5 years cited above, only 2003 is found in the second 25-year interval. As described by Fox et al. (2015) for the Alpes-Maritimes, the decrease in BA corresponds to an improvement in firefighting strategy, since the latter period had some of the hottest summers on record; the same explanation appears to hold for the neighboring administrative divisions studied here.

Figure 9 maps the cumulative percentage area burned inside each 25 ha cell for 1970–1994 and 1995–2019, respectively. Generally, most fires occur in the WUI north of the large coastal cities since densely developed areas have too little vegetation to burn, and relatively remote areas have too few ignition sources. Although we did not treat wind direction or speed, BA shapes in both periods tend to align themselves with known wind patterns in the region. They have a NW–SE orientation throughout most of the western and central sections (Bouches-du-Rhône and Var administrative divisions) but show little preferential orientation in the eastern administrative division of Alpes-Maritimes where wind

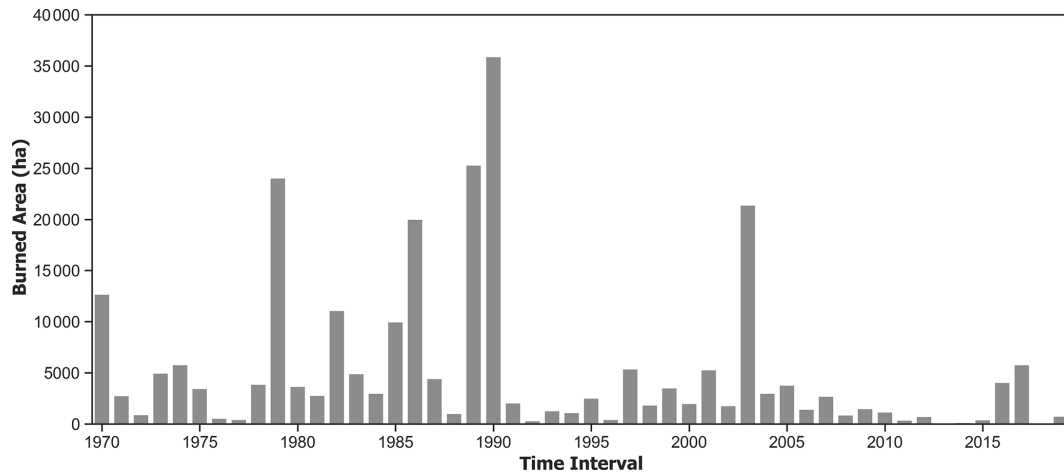


Figure 8. History of annual forested burned areas from 1970 to 2019.

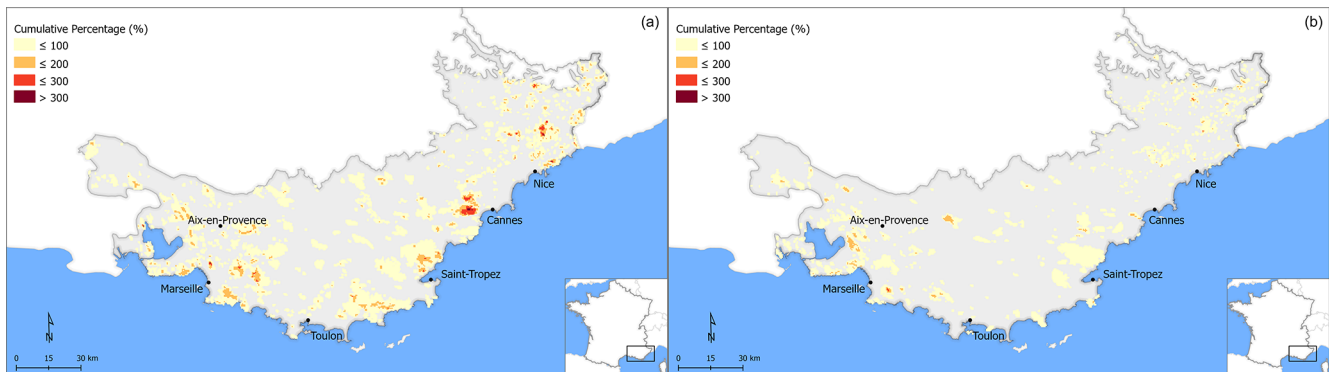


Figure 9. Cumulative percentage of the forested burned areas in the 1970–1994 interval (a) and in the 1995–2019 interval (b) over a 500×500 m grid.

speeds are lower than the mistral winds in the Rhône valley. There is a clear difference between the two periods, with the second one having significantly fewer burned cells which are also slightly more spatially dispersed. In addition, cumulative percentage values are noticeably lower, with a small number of cells (302) exceeding 100 % and very few (9) reaching 200 %. All major hotspots disappear in the second interval, apart from some located mainly in the western area of study zone near Aix-Marseille.

The largest patches in both intervals are found in the central part of the study zone in the Var administrative division, which combines continuous forest cover and a lower population density that is distributed more evenly throughout the administrative division. The two largest continuous BA clusters are found here, with one being north of Saint-Tropez and one east of Toulon. In the 1995–2019 time interval, the first cluster shrunk, whereas the second one completely disappeared. In the western section of the study area (Bouches-du-Rhône), burned patches are located in constrained areas between densely built zones (Aix-en-Provence and Marseille), with several cells displaying high fire recurrence. In the east-

ern section of the study area (Alpes-Maritimes), where population is particularly dense along the coast (Cannes–Nice), BA cells are concentrated inland along the periphery of the coastal built-up area. A major hotspot with the highest cumulative percentage burned area is found just west of Cannes, and this patch almost disappears in the second period. In comparison to the rest of the study area, patches in the eastern administrative division of the Alpes-Maritimes are smaller and more numerous, with high to very high recurrence, even at higher altitudes.

Spatiotemporal analysis

Results of the CMK method depict areas of increasing and decreasing trends in terms of mean annual BA over the study area (Fig. 10). Positive Z scores (colored in red) correspond to areas with increasing trends, and negative Z scores (colored in blue) correspond to areas with decreasing trends. Overall, a general decreasing trend of BA throughout most of the study area can be observed, with approximately 60 % of the cells corresponding to a negative value. The largest

clusters of negative Z scores are located predominately in the central areas of the region, north of Toulon, north of Saint-Tropez, and west of Cannes, with small negative patches northeast of Marseille and north of Nice. Positive Z score clusters are more constrained in terms of size and are generally dispersed. Significant decreasing trends are relatively limited and can be spotted in areas such as to the east of Marseille, west of Cannes, and north of Nice. Significant positive trends are detected in several locations (although limited in area) such as between Aix-en-Provence and Marseille and in the northeastern part (Alpes-Maritimes administrative division) of the study area. Although contrasting negative–positive trends co-exist in close proximity near Marseille and Aix-en-Provence, the greatest speckled pattern is found in the Alpes-Maritimes administrative division, where fires are smaller and more randomly distributed.

3.2 Fire selectivity and topography

Topographic effects studied here include the slope aspect and inclination. Since some areas may have greater BA values simply because a given topographic class is more frequent in the landscape, Jacobs' selectivity index was calculated in order to identify potential classes of aspect and inclination that are preferred by fire between two periods of (i) 1970–1994 and (ii) 1995–2019.

Figure 11 shows fire preference (Jacobs' index > 0) and fire avoidance (Jacobs' index < 0) for the two 25-year periods under study. Between 1970–1994, S-facing slopes have a weak positive median value (0.02), while the others are all negative. Values become increasingly negative in the following order: W (-0.08), E (-0.12), N (-0.18), and flat (-0.38). In the second period (1995–2019), the median fire selectivity of S-facing slopes (0.1) increases and presents a clear difference with other trends which either remain the same (flat) or decrease. N-facing (-0.33) slopes appear to be even less prone to fire in the 1995–2019 interval, and flat surfaces continue to show the greatest aversion to fire.

As for aspect, Fig. 12 shows fire selectivity for each of the two periods based on Jacobs' selectivity index according to slope inclination. Overall, fire is not selective with regards to inclination; in the first period, the gentlest ($\leq 10^\circ$) and steepest ($> 40^\circ$) inclination categories tend to be avoided by fire (values of -0.20 and -0.19 , respectively). In the second period, median fire selectivity for gentlest slopes ($\leq 10^\circ$) show slightly stronger avoidance, shifting from -0.2 to -0.29 , while steepest ($> 40^\circ$) slopes, located mainly in the eastern segment of the study area, exhibit a similar change, shifting from -0.19 to -0.27 . Intermediate slope categories (10 – 40°), which account for a high percentage of BA in the western (Bouches-du-Rhône) and central (Var) study area, do not exhibit any clear fire selectivity pattern.

3.3 Fire selectivity and vegetation type

Forested and semi-natural vegetation is distributed between five categories, of which natural grasslands and sclerophyllous vegetation have the lowest and the highest 50-year average covers, respectively, as the following values show: broadleaved forest (20.6%), coniferous forest (24.1%), mixed forest (19.2%), natural grasslands (11.2%), and sclerophyllous vegetation (24.9%). Over the 50-year study period, mixed and broadleaved forests maintain roughly the same area, whereas conifers present a slight but decreasing trend. Sclerophyllous vegetation expanded in the study area ($\approx 6\%$ increase), becoming the most common type in the last 3 decades. Finally, natural grassland is by far the least common type and shrunk slightly ($\approx 3.5\%$ decrease) over time.

Fire selectivity with regards to vegetation type is presented in Fig. 13. In the first period, three types of vegetation show signs of fire avoidance, namely mixed (-0.28), broadleaved (-0.24), and coniferous (-0.21). Natural grasslands and sclerophyllous vegetation display a weak preference for fire, with median values of 0.09 and 0.05, respectively.

Even though the order changes slightly in the second period, the effects of the fire suppression strategy on vegetation types are more evident than for the topographic factors. On the one hand, all three forest types are more clearly avoided by fire, while, on the other hand, natural grasslands and sclerophyllous vegetation show even stronger fire preference in the second period, shifting from 0.08 to 0.28 and from 0.05 to 0.15, respectively.

3.4 Geographically weighted regression

There is considerable spatiotemporal variability in the strength of the correlation between the BA and environmental variables throughout the study area. The coefficient of determination R^2 values range spatially from 0.00 to 0.68 (slope inclination), depending on the variable and time interval (Table 4). The explanatory power for all values tends to be weak, and topographic factors and sclerophyllous vegetation show the strongest correlations with BA. The remaining vegetation types display a weak fit that is similar in both periods.

Figures 14 and 15 depict local R^2 results of the application of GWR between the percentage of BA and topographic factors. Overall, the highest values are concentrated mainly in the western and central parts (closer to the coastline) of the study area for both slope aspect and inclination. The proportion of variance explained by aspect is slightly greater in the second period, with several cells being in the highest class (0.25–0.35). Despite having a strong local fit in the first period, both distribution and variability changed drastically for slope inclination in the second period.

Figures 16 to 20 display local R^2 results of the application of GWR between the percentage of BA and the percent-

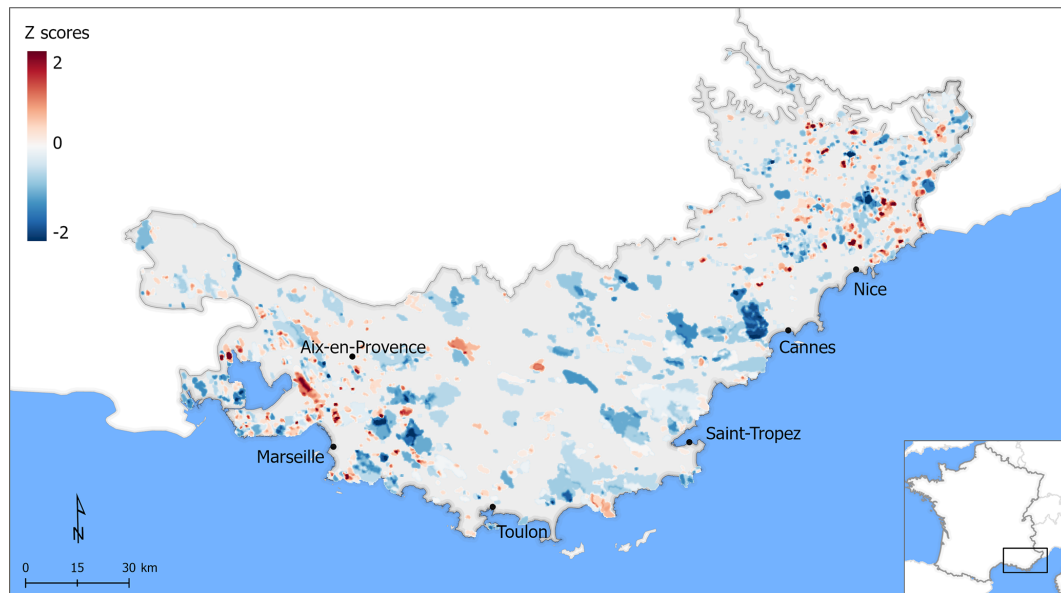


Figure 10. Trends of mean annual burned area between 1970 and 2019, based on the contextual Mann–Kendall method. Areas with positive Z scores depict increasing trends in the burned area, while negative Z scores show decreasing trends.

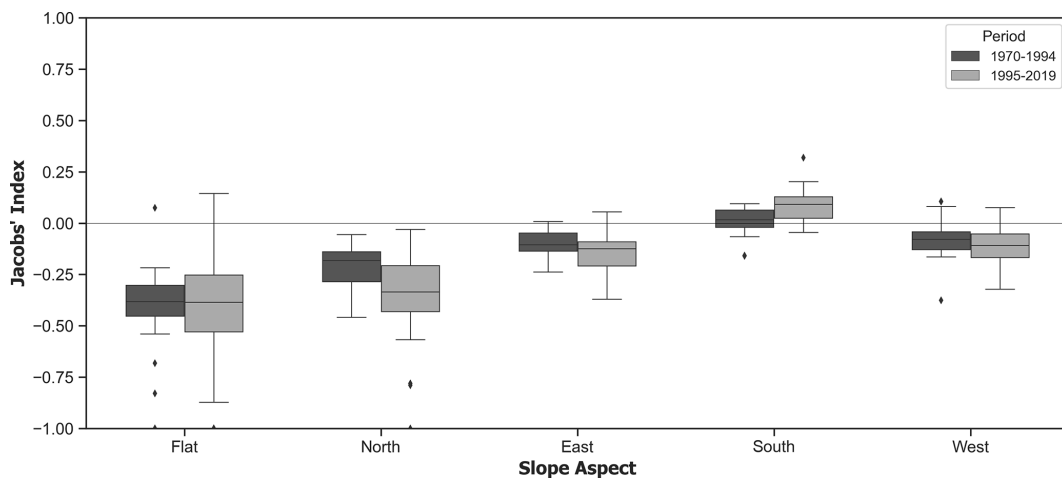


Figure 11. Box plot representing the distribution of the Jacobs' index (ranging from -1 to $+1$) for 1970–1994 (left) and 1995–2019 (right) according to the slope aspect. The (i) median value (50th percentile; bar within the box), (ii) first quartile (25th percentile; bottom part of the box), and (iii) third quartile (75th percentile; top part of the box) are shown. Whiskers represent observations outside the middle 50%, and points represent outliers.

age of each vegetation type. Similar to topographic variables, sclerophyllous vegetation exhibits the same spatial pattern of high R^2 values. A clear increase in local R^2 can be observed when moving towards the western part of the region, which is more evident in the first period. Low fits are found for both periods in the higher altitude areas, located mainly in north-eastern segments of the area. R^2 values for natural grasslands are generally low and display small differences both in terms of space and variance. Explanatory variables related to forest categories show a very weak fit in the relationship with BA.

In addition, the general clustering patterns are quite different between the two periods.

4 Discussion

4.1 Fire history

BA in southeastern France has undergone substantial changes over the last 50 years. Annually, BA varies considerably, but clear declining trends are observed in the second part of the temporal interval under study. Around half of the

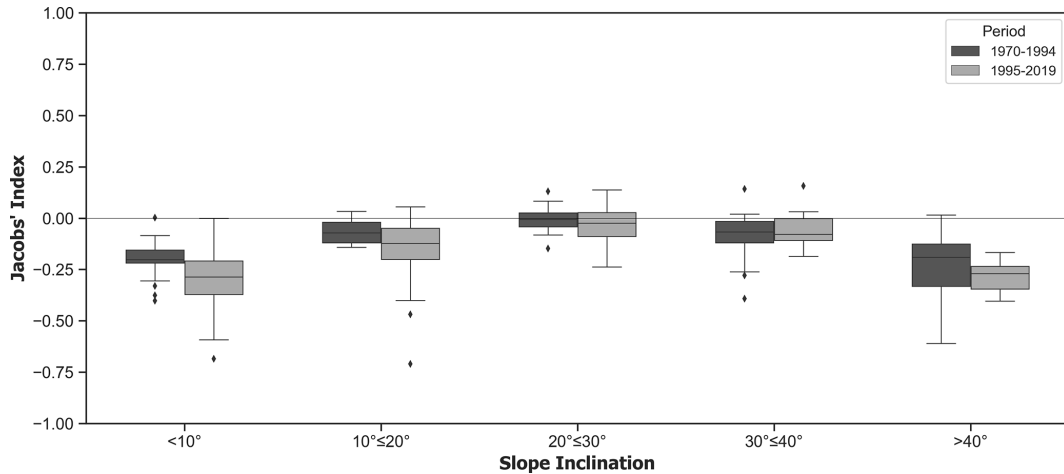


Figure 12. Box plot representing the distribution of the Jacobs' index (ranging from -1 to $+1$) for 1970–1994 (left) and 1995–2019 (right), according to slope inclination. The (i) median value (50th percentile; bar within the box), (ii) first quartile (25th percentile; bottom part of the box), and (iii) third quartile (75th percentile; top part of the box) are shown. Whiskers represent observations outside the middle 50 %, and points represent outliers.

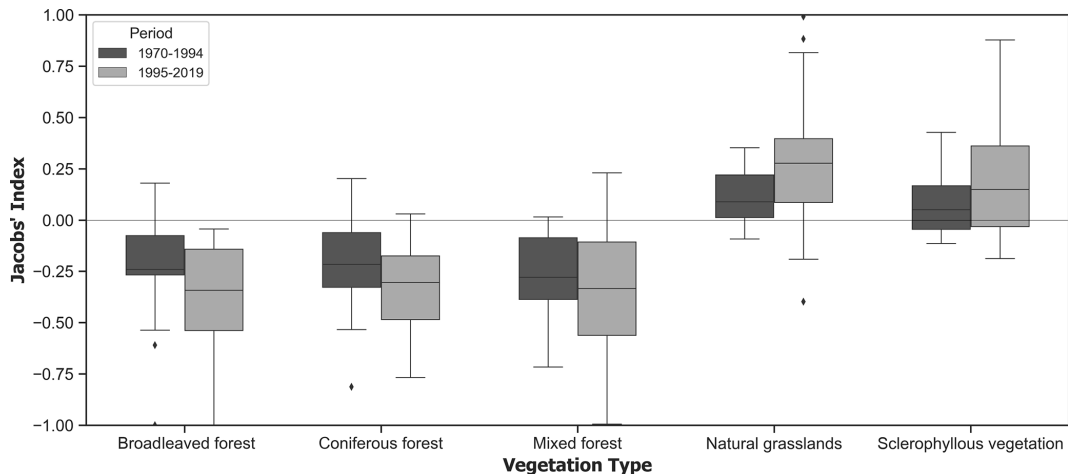


Figure 13. Box plot representing the distribution of the Jacobs' index (ranging from -1 to $+1$) for 1970–1994 (left) and 1995–2019 (right) according to vegetation type. The (i) median value (50th percentile; bar within the box), (ii) first quartile (25th percentile; bottom part of the box), and (iii) third quartile (75th percentile; top part of the box) are shown. Whiskers represent observations outside the middle 50 %, and points represent outliers.

total BA (126,700 ha) was recorded in 5 years, i.e., 1979, 1986, 1989, 1990, and 2003. Due to particularly catastrophic fires in the 1980s, a new fire suppression policy (Vulcain) was initiated that came fully into effect in 1994 (Battesti, 1997). This new strategy focused on aggressively suppressing fire ignitions under any weather conditions in order to avoid fire propagation to the extent where suppression would become both more difficult and more expensive. Although the fire weather index values were not calculated here for the three administrative divisions, Fox et al. (2015) noted a general increase in summer temperatures between about 1980 and 2010, so the firefighting policy had a major impact on the decrease in total BA after 1994. Only 2003 stands out

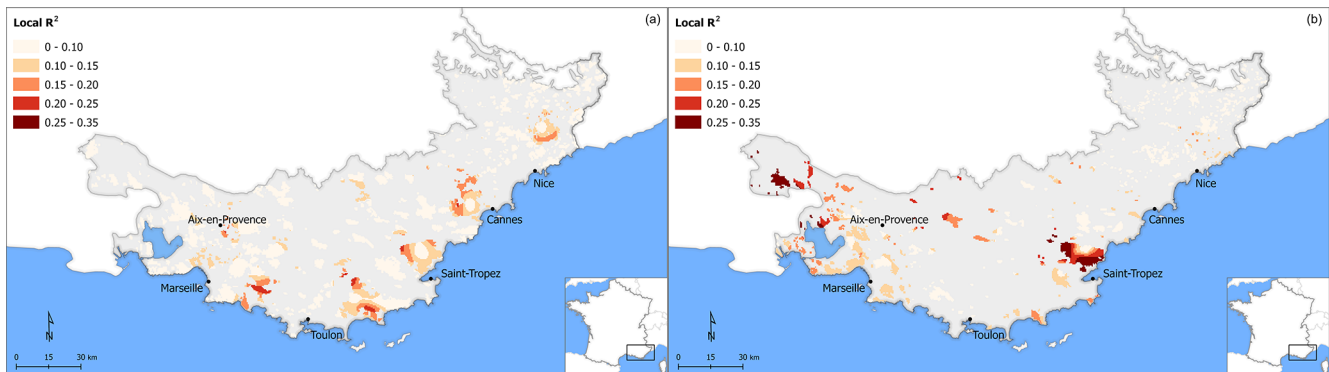
as a big fire year in the 1995–2019 interval, and although it was the hottest/driest year on record in the Alpes-Maritimes, it remained within the range of BA values of the big 1980s fires (Fox et al., 2015). Nonetheless, it raises doubts about the sustainability of rapid suppression in extreme conditions when resources are spread thinly over a greater number of ignitions (Curt and Frejaville, 2018).

Spatiotemporal analysis

The effect of the new firefighting strategy can also be viewed spatially. In general, fire patches are less large and are distributed over smaller geographic proximities with one an-

Table 4. Descriptive statistics of local R^2 per environmental factor for 1970–1994 (P1) and for 1995–2019 (P2).

Period	Slope aspect		Slope inclination		Sclerophyllous vegetation		Natural grasslands		Coniferous forest		Broadleaved forest		Mixed forest		
	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Maximum	0.24	0.36	0.68	0.25	0.48	0.47	0.19	0.21	0.20	0.21	0.11	0.23	0.26	0.25	
Mean	0.08	0.11	0.13	0.06	0.19	0.17	0.07	0.08	0.08	0.09	0.04	0.06	0.07	0.05	
Median	0.07	0.1	0.1	0.05	0.16	0.15	0.08	0.06	0.05	0.03	0.03	0.05	0.04	0.05	
Standard deviation	0.08	0.12	0.08	0.04	0.12	0.11	0.03	0.05	0.03	0.03	0.02	0.05	0.05	0.04	

**Figure 14.** Spatial distribution of local R^2 between the burned area and slope aspect for 1970–1994 (a) and 1995–2019 (b).

other, and the fire recurrence is lower. Spatiotemporal trends, however, vary from west to east, according to the specific population and environmental contexts of each administrative division. In the western part of the study zone, around Aix-en-Provence and Marseille, hotspots, in the form of positive Z scores, remain, and the new firefighting strategy had less effect since fires were already limited in size by vegetation continuity. Although limited in area, multiple clusters of positive trends are found in closer proximity to the built-up areas near Marseille and Aix-en-Provence in comparison to overall decreasing trends. Increased human activity is known to affect fire ignition (Badia et al., 2011; Chas-Amil et al., 2013; Jiménez-Ruano et al., 2017; Lampin-Maillet et al., 2011), and in our context that can be potentially linked to the high number of arson activities found in the area (Curt et al., 2016). On the contrary, in the central part of the study area, where most of the big fires occur, the new fire policy effectively limited fire propagation over the continuous vegetated cover that defines the region. This zone displays the largest clusters of negative Z scores, thereby decreasing BA with very few positive values and low fire recurrence. Ganteaume and Barbero (2019) provided evidence that large fires (> 100 ha) declined sharply in the central segment of the study area after the introduction of the fire management policy, and our results, using different methods, are coherent with their findings. Finally, in the eastern segment of the study area, frequent small dispersed fire patches are found.

Fire shapes are not elongated by wind direction compared to polygons in the western and central administrative divisions, and although negative fire occurrence trends dominate, particularly in the WUI band, there is a greater number of small positive patches compared to other zones.

4.2 Burned area and topography

S-facing slopes have the greatest BA, burn more frequently (Mouillot et al., 2003), and are more exposed to forest fires than other slopes due to both environmental factors (greater insolation and evapotranspiration) and WUI characteristics since S-facing slopes in southern France have more houses and therefore more potential ignition sources (Fox et al., 2018). S-facing (sum of SW, S, and SE) slopes play an increasingly important role over time, and this could be linked to a combination of hotter summers and an increasing number of human dwellings on these slopes, as growth rates on S-facing slopes in the Alpes-Maritimes were 4–5 times greater than on N-facing slopes in 1990–2012.

Slope inclination favors fire propagation directly through more efficient radiative heat transfer (Rothermel, 1983) and increases the rate of spread and fire intensity (Csontos and Cseresnyés, 2015; Capra et al., 2018). In addition, slope inclination influences fire ignition and suppression indirectly through accessibility, solar radiation variations, fuel moisture, and fuel density, which in turn influence flammability

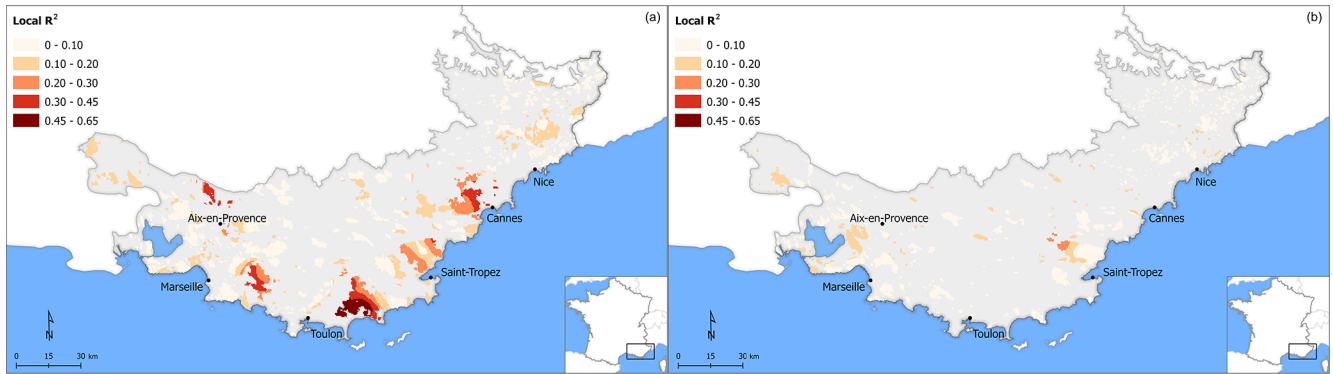


Figure 15. Spatial distribution of local R^2 between the burned area and slope inclination for 1970–1994 (a) and 1995–2019 (b).

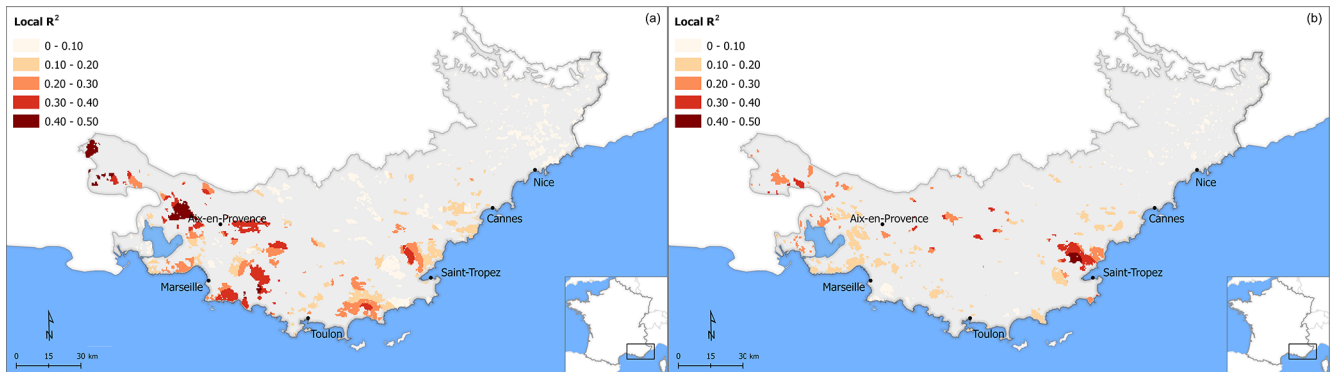


Figure 16. Spatial distribution of local R^2 between the burned area and percent cover of the sclerophyllous vegetation for 1970–1994 (a) and 1995–2019 (b).

(Holden et al., 2009). In this study, flat areas are most avoided by fire for several independent reasons. Radiative heat transfer is less efficient on these slopes, which are more densely inhabited and more easily accessible with denser road networks, so lower fire preference probably depends as much or more on early suppression as on physical processes. The fire avoidance of low slope inclinations strengthens over time, and this is coherent with more rapid suppression in this interval. The BA in intermediate slope inclinations is not affected significantly by the change in firefighting strategy, potentially due to factors that counter rapid suppression like more efficient radiative heat transfer, more difficult accessibility, and the presence of isolated or diffuse housing.

4.3 Burned area and vegetation type

The role of vegetation in fire frequency and BA patches located in the Bouches-du-Rhône and Var administrative divisions was studied by Curt et al. (2013). Their case study reflects patterns observed here at a larger scale, namely that vegetation flammability is secondary to landscape organization. Large open patches of continuous fuel, as are found in the Var administrative division, favor larger fires with longer return intervals than the small patchy wildland distribution

in the Bouches-du-Rhône (Ganteaume and Barbero, 2019). Burned vegetation patterns observed here highlight the frequently cited role of sclerophyllous vegetation (shrubland) (Ganteaume and Jappiot, 2013; Moreira et al., 2011; Oliveira et al., 2014a; Tessler et al., 2016). Shrublands both favor fire propagation in dry conditions (Baeza et al., 2002) and result from recurrent fires (Tessler et al., 2016). As Mermoz et al. (2005) suggested, the fire proneness of sclerophyllous vegetation is connected to its ability to regenerate faster and generate quicker fuel accumulation; this also applies in our case, since sclerophyllous vegetation covers the greatest area, has the greatest BA and greatest explained variance in the GWR analysis, and is one of two vegetation categories (with natural grasslands) that has positive resource index values. These results are coherent with the findings of others working in Mediterranean environments where large fires tend to occur in landscapes with dense shrublands (Moreira et al., 2011; Ruffault and Mouillot, 2017). In a context where initial suppression is crucial to fire extinction, sclerophyllous vegetation may resist early suppression better than other covers, where initial propagation is perhaps slower. Moreover, firefighting assets appear to prioritize other types of vegetation during fire suppression, since fire selectivity remains

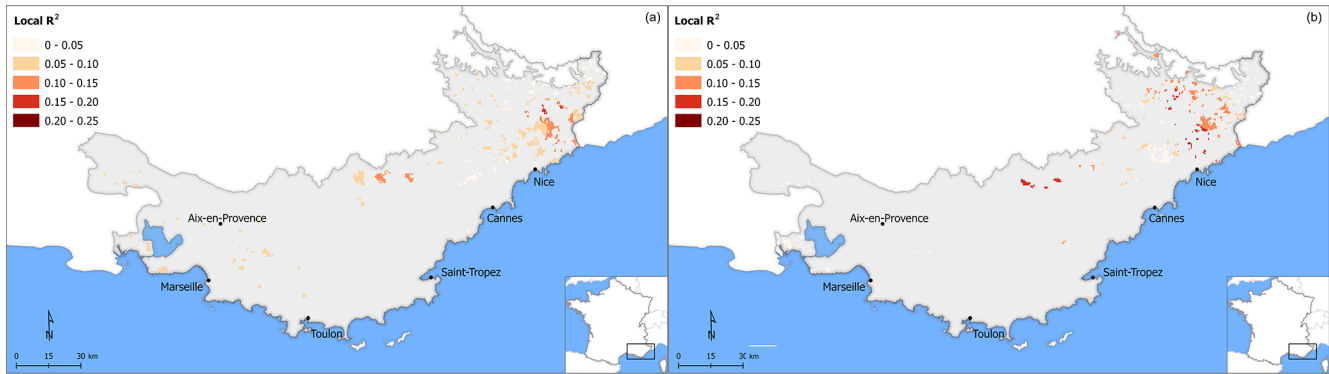


Figure 17. Spatial distribution of local R^2 between burned area and percent cover of natural grasslands for 1970–1994 (a) and 1995–2019 (b).

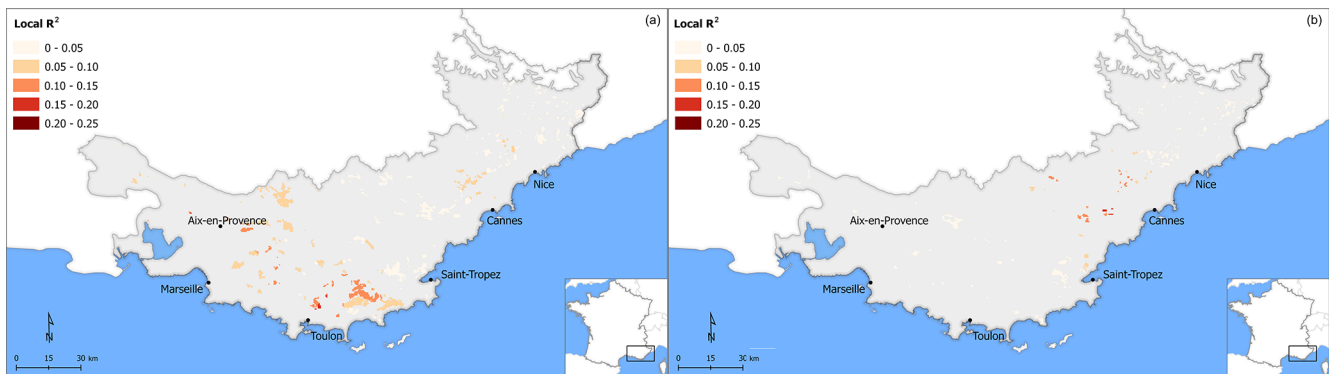


Figure 18. Spatial distribution of local R^2 between the burned area and percent cover of the coniferous forest for 1970–1994 (a) and 1995–2019 (b).

unchanged for bushlands, possibly due to the low cost of restoration (Oehler et al., 2012).

As other studies have concluded (Oliveira et al., 2014a), natural grasslands display a high fire susceptibility. Despite the change in the firefighting policy, grasslands are over-represented in BA in both time intervals, and this may be due to faster initial propagation or accessibility issues as, for example, in certain mid- to high-altitude areas over the eastern section of the study area, where burned clusters of this vegetation type are found. Sheep grazing is a common practice in high alpine pastures of the Alpes-Maritimes administrative division, and natural grassland fires may be due to bush clearing operations by shepherds which resulted in uncontrolled wildfires that affected much larger areas than originally intended. All three forest types (broadleaved, coniferous, and mixed) display a similar pattern characterized by fire avoidance that is even more evident after the fire management policy change. This does not necessarily reflect a higher priority for suppression by firefighting assets over other vegetation types but may indicate that fires in these vegetation types initially take more time to spread than in bushland, so they are suppressed before becoming large fires.

5 Conclusion

In this study, the results provide a coherent picture of the impact of a shift in firefighting strategy on fire occurrence and environmental characteristics. The burned area decreased sharply in the SE of France after 1994 with the introduction of the new firefighting strategy. Rapid fire extinction was particularly effective in limiting big fires in the region. Large fire hotspots found mainly in the central parts disappear after the policy change, while new clusters of high fire recurrence appear in closer proximity to areas with increased human activity.

S-facing aspects have an increasingly bigger impact over time, and this may be linked to both environmental conditions and increased human presence on those slopes. Fire avoids low slope inclinations and even more so after the shift in fire suppression, as flat areas are easier to access and more densely inhabited, so lower fire preference is probably determined as much or more by early suppression as by physical processes (reduced radiative heat transfer).

Over half of the total BA in the last 50 years concerned sclerophyllous vegetation, thus confirming its strong association with high fire susceptibility and recurrence. Considering that sclerophyllous vegetation regenerates and expands

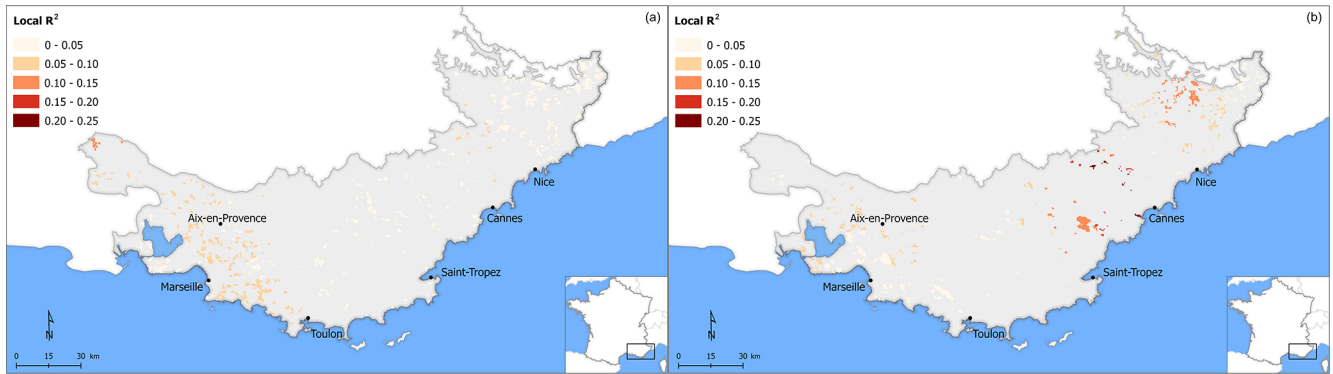


Figure 19. Spatial distribution of local R^2 between the burned area and percent cover of the broadleaved forest for 1970–1994 (a) and 1995–2019 (b).

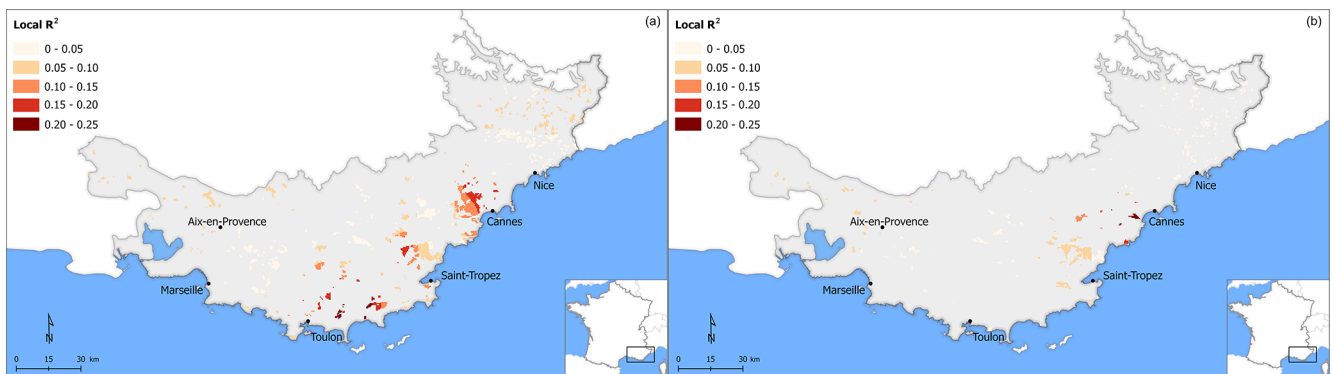


Figure 20. Spatial distribution of local R^2 between the burned area and percent cover of the mixed forest for 1970–1994 (a) and 1995–2019 (b).

faster than other vegetation types in the region, this may lead to an increase in fire risk in the future. Natural grasslands, even though they cover limited area and decline with time, are also preferred by fire which may be due to pastoral fires. On the contrary, broadleaved, coniferous, and mixed forest are avoided by fire, especially after the change in fire management policy.

Further ongoing exploitation of the fire GIS database in conjunction with WUI characteristics will likely further improve our understanding on the driving forces of BA and the impacts of firefighting strategies in the region.

Code availability. Python packages seaborn (Waskom, 2021), pandas (McKinney, 2010), and matplotlib (Hunter, 2007) were used to process the data and produce the figures. The code is available upon request from the corresponding author.

Data availability. The use of the fire database is subject to specific authorization from the National Forestry Office (Office National des Forêts – ONF) and the Delegation for the Protection of

the Mediterranean Forest (Délégation à la Protection de la Forêt Méditerranéenne – DPFM).

Author contributions. CB established the fire geodatabase, carried out the data processing, analyses, visualization, and wrote the initial draft. DMF performed the land cover modeling, contributed to the interpretation of the results, and reviewed the paper. EDB provided expertise for the data analyses and reviewed the paper.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors wish to thank the National Forestry Office (Office National des Forêts, ONF), for providing the fire geo-database. We would also like to thank the two reviewers, for

their constructive feedback that helped to improve the quality of the paper.

Financial support. This research has been supported by the French government through the 3IA Côte d’Azur and UCAJEDI Investments in the Future project managed by the National Research Agency (grant nos. ANR-19-P3IA-0002 and ANR-15-IDEX-01).

Review statement. This paper was edited by Ricardo Trigo and reviewed by Julien Ruffault and Olga Viedma.

References

- Akaike, H.: A new look at the statistical model identification, *IEEE Trans. Automat. Contr.*, 19, 716–723, <https://doi.org/10.1109/TAC.1974.1100705>, 1974.
- Alexander, J. D., Seavy, N. E., Ralph, C. J., and Hogo-boom, B.: Vegetation and topographical correlates of fire severity from two fires in the Klamath-Siskiyou region of Oregon and California, *Int. J. Wildl. Fire*, 15, 237–245, <https://doi.org/10.1071/WF05053>, 2006.
- Badia, A., Serra, P., and Modugno, S.: Identifying dynamics of fire ignition probabilities in two representative Mediterranean wildland-urban interface areas, *Appl. Geogr.*, 31, 930–940, <https://doi.org/10.1016/j.apgeog.2011.01.016>, 2011.
- Baeza, M. J., De Luís, M., Raventós, J., and Escarré, A.: Factors influencing fire behaviour in shrublands of different stand ages and the implications for using prescribed burning to reduce wildfire risk, *J. Environ. Manage.*, 65, 199–208, <https://doi.org/10.1006/jema.2002.0545>, 2002.
- Bajocco, S. and Ricotta, C.: Evidence of selective burning in Sardinia (Italy): Which land-cover classes do wildfires prefer?, *Landsc. Ecol.*, 23, 241–248, <https://doi.org/10.1007/s10980-007-9176-5>, 2008.
- Barros, A. M. G. and Pereira, J. M. C.: Wildfire selectivity for land cover type: Does size matter?, *PLoS One*, 9, e84760, <https://doi.org/10.1371/journal.pone.0084760>, 2014.
- Battesti, A.: Vulcain ou l’histoire d’une tentative pour commencer à changer les choses dans le système “Feux de Forêts”, par la transparence et la convivialité, *Forêt Méditerranéenne*, XVIII, 143–145, hal-03558006, 1997.
- Birch, D. S., Morgan, P., Kolden, C. A., Abatzoglou, J. T., Dillon, G. K., Hudak, A. T., and Smith, A. M. S.: Vegetation, topography and daily weather influenced burn severity in central Idaho and western Montana forests, *Ecosphere*, 6, 1–23, <https://doi.org/10.1890/ES14-00213.1>, 2015.
- Bond, W. J. and Keeley, J. E.: Fire as a global “herbivore”: the ecology and evolution of flammable ecosystems, *Trends Ecol. Evol.*, 20, 387–394, 2005.
- Bowman, D. M. J. S., Williamson, G. J., Abatzoglou, J. T., Kolden, C. A., Cochrane, M. A., and Smith, A. M. S.: Human exposure and sensitivity to globally extreme wildfire events, *Nat. Ecol. Evol.*, 1, 0058, <https://doi.org/10.1038/s41559-016-0058>, 2017.
- Broncano, M. J. and Retana, J.: Topography and forest composition affecting the variability in fire severity and post-fire regeneration occurring after a large fire in the Mediterranean basin, *Int. J. Wildl. Fire*, 13, 209–216, <https://doi.org/10.1071/WF03036>, 2004.
- Capra, G. F., Tidu, S., Lovreglio, R., Certini, G., Salis, M., Bacciu, V., Ganga, A., and Filzmoser, P.: The impact of wildland fires on calcareous Mediterranean pedosystems (Sardinia, Italy) – An integrated multiple approach, *Sci. Total Environ.*, 624, 1152–1162, <https://doi.org/10.1016/j.scitotenv.2017.12.099>, 2018.
- Carmo, M., Moreira, F., Casimiro, P., and Vaz, P.: Land use and topography influences on wildfire occurrence in northern Portugal, *Landsc. Urban Plan.*, 100, 169–176, <https://doi.org/10.1016/j.landurbplan.2010.11.017>, 2011.
- Catarino, S., Romeiras, M. M., Figueira, R., Aubard, V., Silva, J. M. N., and Pereira, J. M. C.: Spatial and temporal trends of burnt area in angola: Implications for natural vegetation and protected area management, *Diversity*, 12, 307, <https://doi.org/10.3390/D12080307>, 2020.
- Chas-Amil, M. L., Touza, J., and García-Martínez, E.: Forest fires in the wildland-urban interface: A spatial analysis of forest fragmentation and human impacts, *Appl. Geogr.*, 43, 127–137, <https://doi.org/10.1016/j.apgeog.2013.06.010>, 2013.
- Csontos, P. and Cseresnyés, I.: Fire-risk evaluation of austrian pine stands in Hungary - Effects of drought conditions and slope aspect on fire spread and fire behaviour, *Carpathian J. Earth Environ. Sci.*, 10, 247–254, 2015.
- Curt, T. and Frejaville, T.: Wildfire Policy in Mediterranean France: How Far is it Efficient and Sustainable?, *Risk Anal.*, 38, 472–488, <https://doi.org/10.1111/risa.12855>, 2018.
- Curt, T., Borgniet, L., and Bouillon, C.: Wildfire frequency varies with the size and shape of fuel types in southeastern France: Implications for environmental management, *J. Environ. Manage.*, 117, 150–161, <https://doi.org/10.1016/j.jenvman.2012.12.006>, 2013.
- Curt, T., Fréjaville, T., and Lahaye, S.: Modelling the spatial patterns of ignition causes and fire regime features in southern France: Implications for fire prevention policy, *Int. J. Wildl. Fire*, 25, 785–796, <https://doi.org/10.1071/WF15205>, 2016.
- Dickson, B. G., Prather, J. W., Xu, Y., Hampton, H. M., Aumack, E. N., and Sisk, T. D.: Mapping the probability of large fire occurrence in northern Arizona, USA, *Landsc. Ecol.*, 21, 747–761, <https://doi.org/10.1007/s10980-005-5475-x>, 2006.
- Douglas, E. M., Vogel, R. M., and Kroll, C. N.: Trends in floods and low flows in the United States: Impact of spatial correlation, *J. Hydrol.*, 240, 90–105, [https://doi.org/10.1016/S0022-1694\(00\)00336-X](https://doi.org/10.1016/S0022-1694(00)00336-X), 2000.
- Duane, A., Piqué, M., Castellnou, M., and Brotons, L.: Predictive modelling of fire occurrences from different fire spread patterns in Mediterranean landscapes, *Int. J. Wildl. Fire*, 24, 407–418, <https://doi.org/10.1071/WF14040>, 2015.
- Durbin, J. and Watson, G. S.: Testing for Serial Correlation in Least Squares Regression: I, *Biometrika*, 37, 409–428, <https://doi.org/10.2307/2332391>, 1950.
- Eastman, J., R.: TerrSet 2020 Help System, Accessed in TerrSet version 19, Clark University, 394 pp., 2020.
- Elia, M., Giannico, V., Laforteza, R., and Sanesi, G.: Modeling fire ignition patterns in Mediterranean urban interfaces, *Stoch. Environ. Res. Risk Assess.*, 33, 169–181, <https://doi.org/10.1007/s00477-018-1558-5>, 2019.
- Estes, B. L., Knapp, E. E., Skinner, C. N., Miller, J. D., and Preisler, H. K.: Factors influencing fire severity under moderate burning

- conditions in the Klamath Mountains, northern California, USA, *Ecosphere*, 8, e01794, <https://doi.org/10.1002/ecs2.1794>, 2017.
- Evin, G., Curt, T., and Eckert, N.: Has fire policy decreased the return period of the largest wildfire events in France? A Bayesian assessment based on extreme value theory, *Nat. Hazards Earth Syst. Sci.*, 18, 2641–2651, <https://doi.org/10.5194/nhess-18-2641-2018>, 2018.
- Fernandes, P. M., Monteiro-Henriques, T., Guiomar, N., Loureiro, C., and Barros, A. M. G.: Bottom-Up Variables Govern Large-Fire Size in Portugal, *Ecosystems*, 19, 1362–1375, <https://doi.org/10.1007/s10021-016-0010-2>, 2016.
- Fotheringham, A. S., Charlton, M. E., and Brunsdon, C.: Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis, *Environ. Plan. A*, 30, 1905–1927, <https://doi.org/10.1068/a301905>, 1998.
- Fotheringham, A. S., Brunsdon, C., and Martin, C.: *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, John Wiley & Sons, ISBN 0-471-49616-2, 2003.
- Fox, D. M., Martin, N., Carrega, P., Andrieu, J., Adnès, C., Emsellem, K., Ganga, O., Moebius, F., Tortorollo, N., and 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.*, 56, 1–12, <https://doi.org/10.1016/j.apgeog.2014.10.001>, 2015.
- Fox, D. M., Carrega, P., Ren, Y., Caillouet, P., Bouillon, C., and Robert, S.: How wildfire risk is related to urban planning and Fire Weather Index in SE France (1990–2013), *Sci. Total Environ.*, 621, 120–129, <https://doi.org/10.1016/J.SCITOTENV.2017.11.174>, 2018.
- Ganteaume, A. and Barbero, R.: Contrasting large fire activity in the French Mediterranean, *Nat. Hazards Earth Syst. Sci.*, 19, 1055–1066, <https://doi.org/10.5194/nhess-19-1055-2019>, 2019.
- Ganteaume, A. and Jappiot, M.: What causes large fires in Southern France, *For. Ecol. Manage.*, 294, 76–85, <https://doi.org/10.1016/j.foreco.2012.06.055>, 2013.
- Ganteaume, A., Camia, A., Jappiot, M., San-Miguel-Ayanz, J., Long-Fournel, M., and Lampin, C.: A review of the main driving factors of forest fire ignition over Europe, *Environ. Manage.*, 51, 651–662, <https://doi.org/10.1007/s00267-012-9961-z>, 2013.
- Holden, Z. A., Morgan, P., and Evans, J. S.: A predictive model of burn severity based on 20-year satellite-inferred burn severity data in a large southwestern US wilderness area, *For. Ecol. Manage.*, 258, 2399–2406, <https://doi.org/10.1016/j.foreco.2009.08.017>, 2009.
- Hunter, J. D.: Matplotlib: A 2D Graphics Environment, *Comput. Sci. Eng.*, 9, 90–95, <https://doi.org/10.1109/MCSE.2007.55>, 2007.
- Jacobs, J.: Quantitative Measurement of Food Selection, *Oecologia*, 14, 413–417, <https://doi.org/10.1385/1-59259-055-1:51>, 1974.
- Jiménez-Ruano, A., Rodríguez Mimbrenero, M., and de la Riva Fernández, J.: Exploring spatial–temporal dynamics of fire regime features in mainland Spain, *Nat. Hazards Earth Syst. Sci.*, 17, 1697–1711, <https://doi.org/10.5194/nhess-17-1697-2017>, 2017.
- Kendall, M. G.: *Rank Correlation Methods*, 4th edn., Griffin, London, ISBN 9780852641996, 1975.
- Kolaneck, A., Szymanowski, M., and Raczek, A.: Human Activity Affects Forest Fires: The Impact of Anthropogenic Factors on the Density of Forest Fires in Poland, *Forests*, 12, 728, <https://doi.org/10.3390/f12060728>, 2021.
- Koutsias, N., Martínez-Fernández, J., and Allgöwer, B.: Do factors causing wildfires vary in space? evidence from geographically weighted regression, *GIScience Remote Sens.*, 47, 221–240, <https://doi.org/10.2747/1548-1603.47.2.221>, 2010.
- Lampin-Maillet, C., Long-Fournel, M., Ganteaume, A., Jappiot, M., and Ferrier, J. P.: Land cover analysis in wildland-urban interfaces according to wildfire risk: A case study in the South of France, *For. Ecol. Manage.*, 261, 2200–2213, <https://doi.org/10.1016/j.foreco.2010.11.022>, 2011.
- Manly, B., McDonald, L., Thomas, D., McDonald, T., and Erickson, W.: *Resource Selection by Animals Statistical Design and Analysis for Field Studies*, Second Edition, Kluwer Acad. Publ., ISBN 1-4020-0677-2, 2002.
- Mann, H. B.: Nonparametric Tests Against Trend, *Econometrica*, 13, 245–259, <https://doi.org/10.2307/1907187>, 1945.
- Martínez-Fernández, J., Chuvieco, E., and Koutsias, N.: Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression, *Nat. Hazards Earth Syst. Sci.*, 13, 311–327, <https://doi.org/10.5194/nhess-13-311-2013>, 2013.
- McKinney, W.: Data Structures for Statistical Computing in Python, in: *Proceedings of the 9th Python in Science Conference*, Austin, Texas, USA, 28 June–3 July 2010, edited by: van der Walt, S. and Millman, J., 56–61, <https://doi.org/10.25080/Majora-92bf1922-00a>, 2010.
- Mermoz, M., Kitzberger, T., and Veblen, T. T.: Landscape influences on occurrence and spread of wildfires in Patagonian forests and shrublands, *Ecology*, 86, 2705–2715, <https://doi.org/10.1890/04-1850>, 2005.
- Mhawe, M., Faour, G., and Adjizian-Gerard, J.: Wildfire Likelihood’s Elements: A Literature Review, *Challenges*, 6, 282–293, <https://doi.org/10.3390/challe6020282>, 2015.
- Michelaki, C., Fyllas, N. M., Galanidis, A., Aloupi, M., Evangelou, E., Arianoutsou, M., and Dimitrakopoulos, P. G.: Adaptive flammability syndromes in thermo-Mediterranean vegetation, captured by alternative resource-use strategies, *Sci. Total Environ.*, 718, 137437, <https://doi.org/10.1016/j.scitotenv.2020.137437>, 2020.
- Miller, J. D., Safford, H. D., Crimmins, M., and Thode, A. E.: Quantitative Evidence for Increasing Forest Fire Severity in the Sierra Nevada and Southern Cascade Mountains, California and Nevada, USA, *Ecosystems*, 12, 16–32, <https://doi.org/10.1007/s10021-008-9201-9>, 2009.
- Molina, J. R., Martín, T., Rodríguez Y Silva, F., and Herrera, M. Á.: The ignition index based on flammability of vegetation improves planning in the wildland-urban interface: A case study in Southern Spain, *Landsc. Urban Plan.*, 158, 129–138, <https://doi.org/10.1016/j.landurbplan.2016.11.003>, 2017.
- Molina-Terrén, D. M., Xanthopoulos, G., Diakakis, M., Ribeiro, L., Caballero, D., Delogu, G. M., Viegas, D. X., Silva, C. A., and Cardil, A.: Analysis of forest fire fatalities in Southern Europe: Spain, Portugal, Greece and Sardinia (Italy), *Int. J. Wildl. Fire*, 28, 85–98, <https://doi.org/10.1071/WF18004>, 2019.
- Moreira, F., Rego, F. C., and Ferreira, P. G.: Temporal (1958–1995) pattern of change in a cultural landscape of northwestern Portugal: Implications for fire occurrence, *Landsc. Ecol.*, 16, 557–567, <https://doi.org/10.1023/A:1013130528470>, 2001.
- Moreira, F., Vaz, P., Catry, F., and Silva, J. S.: Regional variations in wildfire susceptibility of land-cover types in Portugal: impli-

- cations for landscape management to minimize fire hazard, *Int. J. Wildl. Fire*, 18, 563–574, <https://doi.org/10.1071/WF07098>, 2009.
- Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, E., Barbati, A., Corona, P., Vaz, P., Xanthopoulos, G., Mouillot, F., and Bilgili, E.: Landscape – wildfire interactions in southern Europe: Implications for landscape management, *J. Environ. Manage.*, 92, 2389–2402, <https://doi.org/10.1016/j.jenvman.2011.06.028>, 2011.
- Moreno, J. M., Viedma, O., Zavala, G., and Luna, B.: Landscape variables influencing forest fires in central Spain, *Int. J. Wildl. Fire*, 20, 678–689, <https://doi.org/10.1071/WF10005>, 2011.
- Mouillot, F., Ratte, J. P., Joffre, R., Moreno, J. M., and Rambal, S.: Some determinants of the spatio-temporal fire cycle in a Mediterranean landscape (Corsica, France), *Landsc. Ecol.*, 18, 665–674, <https://doi.org/10.1023/B:LAND.0000004182.22525.a9>, 2003.
- Narayanaraj, G. and Wimberly, M. C.: Influences of forest roads on the spatial patterns of human- and lightning-caused wildfire ignitions, *Appl. Geogr.*, 32, 878–888, <https://doi.org/10.1016/j.apgeog.2011.09.004>, 2012.
- Neeti, N. and Eastman, J. R.: A Contextual Mann-Kendall Approach for the Assessment of Trend Significance in Image Time Series, *Trans. GIS*, 15, 599–611, <https://doi.org/10.1111/j.1467-9671.2011.01280.x>, 2011.
- Nunes, A. N., Lourenço, L., and Meira, A. C. C.: Exploring spatial patterns and drivers of forest fires in Portugal (1980–2014), *Sci. Total Environ.*, 573, 1190–1202, <https://doi.org/10.1016/j.scitotenv.2016.03.121>, 2016.
- Nunes, M. C. S., Vasconcelos, M. J., Pereira, J. M. C., Dasgupta, N., Alldredge, R. J., and Rego, F. C.: Land Cover Type and Fire in Portugal: Do Fires Burn Land Cover Selectively?, *Landsc. Ecol.*, 20, 661–673, <https://doi.org/10.1007/s10980-005-0070-8>, 2005.
- Oehler, F., Oliveira, S., Barredo, J., Camia, A., San-Miguel-Ayanz, J., Pettenella, D., and Mavsar, R.: Assessing European wild fire vulnerability Assessing European wild fire vulnerability, *Geophysical Research Abstracts*, EGU General Assembly, 22–27 April 2012, Vienna, Austria, 14, EGU2012-9452, 2012.
- Oliveira, S., Moreira, F., Boca, R., San-Miguel-Ayanz, J., and Pereira, J. M. C.: Assessment of fire selectivity in relation to land cover and topography: A comparison between Southern European countries, *Int. J. Wildl. Fire*, 23, 620–630, <https://doi.org/10.1071/WF12053>, 2014a.
- Oliveira, S., Pereira, J. M. C., San-Miguel-Ayanz, J., and Lourenço, L.: Exploring the spatial patterns of fire density in Southern Europe using Geographically Weighted Regression, *Appl. Geogr.*, 51, 143–157, <https://doi.org/10.1016/j.apgeog.2014.04.002>, 2014b.
- Oliveras, I., Gracia, M., Moñ, G., and Retana, J.: Factors influencing the pattern of fire severities in a large wildfire under extreme meteorological conditions in the Mediterranean basin, *Int. J. Wildl. Fire*, 18, 755–764, <https://doi.org/10.1071/WF08070>, 2009.
- Otón, G., Pereira, J. M. C., Silva, J. M. N., and Chuvieco, E.: Analysis of trends in the firecci global long term burned area product (1982–2018), *Fire*, 4, 74, <https://doi.org/10.3390/fire4040074>, 2021.
- Padilla, M. and Vega-García, C.: On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain, *Int. J. Wildl. Fire*, 20, 46–58, <https://doi.org/10.1071/WF09139>, 2011.
- Parks, S. A., Holsinger, L. M., Panunto, M. H., Jolly, W. M., Dobrowski, S. Z., and Dillon, G. K.: High-severity fire: evaluating its key drivers and mapping its probability across western US forests, *Environ. Res. Lett.*, 13, 044037, <https://doi.org/10.1088/1748-9326/aab791>, 2018.
- Pereira, M. G., Aranha, J., and Amraoui, M.: Land cover fire proneness in Europe, *For. Syst.*, 23, 598–610, <https://doi.org/10.5424/fs/2014233-06115>, 2014.
- Pokorná, L., Kučerová, M., and Huth, R.: Annual cycle of temperature trends in Europe, 1961–2000, *Glob. Planet. Change*, 170, 146–162, <https://doi.org/10.1016/j.gloplacha.2018.08.015>, 2018.
- Rodrigues, M., Jiménez, A., and de la Riva, J.: Analysis of recent spatial-temporal evolution of human driving factors of wildfires in Spain, *Nat. Hazards*, 84, 2049–2070, <https://doi.org/10.1007/s11069-016-2533-4>, 2016.
- Rodrigues, M., Jiménez-Ruano, A., and de la Riva, J.: Fire regime dynamics in mainland Spain. Part 1: Drivers of change, *Sci. Total Environ.*, 721, 135841, <https://doi.org/10.1016/j.scitotenv.2019.135841>, 2020.
- Rothermel, R. C.: How to predict the spread and intensity of forest and range fires, *US Dep. Agric. For. Serv. Gen. Tech. Rep.*, INT-143, <https://doi.org/10.2737/INT-GTR-143>, 1983.
- Roy, H. G., Fox, D. M., and Emsellem, K.: Spatial dynamics of land cover change in a Euro-Mediterranean catchment (1950–2008), *J. Land Use Sci.*, 10, 277–297, <https://doi.org/10.1080/1747423X.2014.898105>, 2015.
- Ruffault, J. and Mouillot, F.: How a new fire-suppression policy can abruptly reshape the fire-weather relationship, *Ecosphere*, 6, 1–19, <https://doi.org/10.1890/ES15-00182.1>, 2015.
- Ruffault, J. and Mouillot, F.: Contribution of human and biophysical factors to the spatial distribution of forest fire ignitions and large wildfires in a French Mediterranean region, *Int. J. Wildl. Fire*, 26, 498–508, <https://doi.org/10.1071/WF16181>, 2017.
- San-Miguel-Ayanz, J., Moreno, J. M., and Camia, A.: Analysis of large fires in European Mediterranean landscapes: Lessons learned and perspectives, *For. Ecol. Manage.*, 294, 11–22, <https://doi.org/10.1016/j.foreco.2012.10.050>, 2013.
- San-Miguel-Ayanz, J., Durrant, T., Boca, R., Maianti, P., Libertà, G., Artes Vivancos, T., Jacome Felix Oom, D., Branco, A., De Rigo, D., Ferrari, D., Pfeiffer, H., Grecchi, R., Nuijten, D., and Leray, T.: *Forest Fires in Europe, Middle East and North Africa 2019*, Publications Office of the European Union, Luxembourg, <https://doi.org/10.2760/893>, 2020.
- Silva, J. M. N., Moreno, M. V., Le Page, Y., Oom, D., Bistinas, I., and Pereira, J. M. C.: Spatiotemporal trends of area burnt in the Iberian Peninsula, 1975–2013, *Reg. Environ. Chang.*, 19, 515–527, <https://doi.org/10.1007/s10113-018-1415-6>, 2019.
- Tessler, N., Wittenberg, L., and Greenbaum, N.: Vegetation cover and species richness after recurrent forest fires in the Eastern Mediterranean ecosystem of Mount Carmel, Israel, *Sci. Total Environ.*, 572, 1395–1402, <https://doi.org/10.1016/j.scitotenv.2016.02.113>, 2016.
- Tobler, W. R.: A Computer Movie Simulating Urban Growth in the Detroit Region, *Econ. Geogr.*, 46, 234–240, <https://doi.org/10.2307/143141>, 1970.
- Turco, M., Bedia, J., Di Liberto, F., Fiorucci, P., Von Hardenberg, J., Koutsias, N., Llasat, M. C., Xystrakis, F., and Provenzale,

- A.: Decreasing fires in mediterranean Europe, *PLoS One*, 11, e0150663, <https://doi.org/10.1371/journal.pone.0150663>, 2016.
- Urbieto, I. R., Franquesa, M., Viedma, O., and Moreno, J. M.: Fire activity and burned forest lands decreased during the last three decades in Spain, *Ann. For. Sci.*, 76, 90, <https://doi.org/10.1007/s13595-019-0874-3>, 2019.
- Viedma, O., Urbieto, I. R., and Moreno, J. M.: Wildfires and the role of their drivers are changing over time in a large rural area of west-central Spain, *Sci. Rep.*, 8, 17797, <https://doi.org/10.1038/s41598-018-36134-4>, 2018.
- Wang, X. L. and Swail, V. R.: Changes of extreme Wave Heights in northern Hemisphere Oceans and related atmospheric circulation regimes, *J. Climate*, 14, 2204–2221, [https://doi.org/10.1175/1520-0442\(2001\)014<2204:COEWHI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<2204:COEWHI>2.0.CO;2), 2001.
- Waskom, M. L.: seaborn: statistical data visualization, *Journal of Open Source Software*, 6, 3021, <https://doi.org/10.21105/joss.03021>, 2021.

ADVANCES IN FOREST FIRE RESEARCH

2022

Edited by
**DOMINGOS XAVIER VIEGAS
LUÍS MÁRIO RIBEIRO**

Fire cause classification of undetermined fires in southeastern France

Christos Bountzouklis*¹; Dennis Michael Fox¹; Elena Di Bernardino²

¹*Univerisity of Côte d'Azur, UMR CNRS 7300 ESPACE, Nice, 06204, France, {christos.bountzouklis@univ-cotedazur.fr; dennis.fox@univ-cotedazur.fr}*

²*Univerisity of Côte d'Azur, UMR CNRS 7351 LJAD, Nice, 06108, France, {elena.di_bernardino@univ-cotedazur.fr}*

**Corresponding author*

Keywords

ignition causes, machine learning, arson fires, southeastern France

Abstract

Knowledge about fire ignition causes and their spatiotemporal patterns can greatly enhance the efficiency of fire management and fire strategies. In France, the majority of forest fire research is based on a 2x2 km gridded database that provides amongst other information, the cause of fire ignition. According to the same database however, approximately 70% of all fires between 1973 and 2020 were recorded without a cause of ignition. Therefore, information on fire causes for a very large part of the fires that were recorded in the last 50 years is not taken into consideration and can potentially provide significant evidence on fire ignition patterns. As arson fires are of particular interest in southeastern France since they are the most frequent and account for the largest volume of burned area, this study aimed to exploit a fire ignition point geodatabase and machine learning methodologies, to create and evaluate a model that can identify whether unknown caused fires can be classified as arson or non-arson fires, based on numerous environmental and anthropogenic factors. The results of the study suggest that cause identification can be adequately accurate using such a model, although a larger fire database would increase overall performance.

1. Introduction

Fire ignition patterns can vary significantly both temporally and spatially depending on the cause of ignition (Curt et al., 2016) and can be impacted by a plethora of environmental and anthropogenic drivers (Catry et al., 2009; Syphard et al., 2008; Syphard and Keeley, 2015). Some studies have demonstrated that arson fires can potentially be predicted both spatially and temporally (Gonzalez-Olabarria et al., 2012; Penman et al., 2013). In SE France, arson (particularly pyromania and conflict/interest) is the most frequent ignition cause for large fires (100>ha) (Ganteaume and Jappiot, 2013). Recording fire causes and studying their spatiotemporal patterns is important for establishing useful fire policies (Rodrigues et al., 2014) since a better understanding can enhance the efficacy of fire prevention measures (Oliveira et al., 2012). In France according to the national fire database (Prométhée) that contains records of fire ignition causes, approximately 70% of all fires between 1973 and 2020 were ignited by an unknown cause. The percentage of non-identified causes is high, and it is an additional constraint to the already limited research conducted on fire ignition causes. As in many other disciplines, applications of machine learning methodologies have seen a significant increase in wildfire science over the past years (Jain et al., 2020). Thus, this study aims to examine whether a fire ignition point dataset coupled with machine learning methods can be used to identify the source (arson or non-arson) of unknown caused fires and evaluate the importance of several environmental and anthropogenic factors in determining the ignition source.

2. Data & Methodology

2.1. Study area

The study area covers the administrative department of “Bouches-du-Rhône”, which according to the official forest fire database in France (Prométhée), ranks second in terms of burned area and fire frequency in mainland France (Figure 1, Table 1). The department is characterized by gentle slopes and low to intermediate altitudes that increase when moving eastwards. Population density (388.8 people/km²) is higher in the eastern half of the department since that is where the second most populated city in France (Marseille) is found and because the

westernmost parts are covered by wetlands and a national park. Therefore, the westernmost section has a low potential for fire ignition and propagation but increases when moving towards the eastern half of the department.

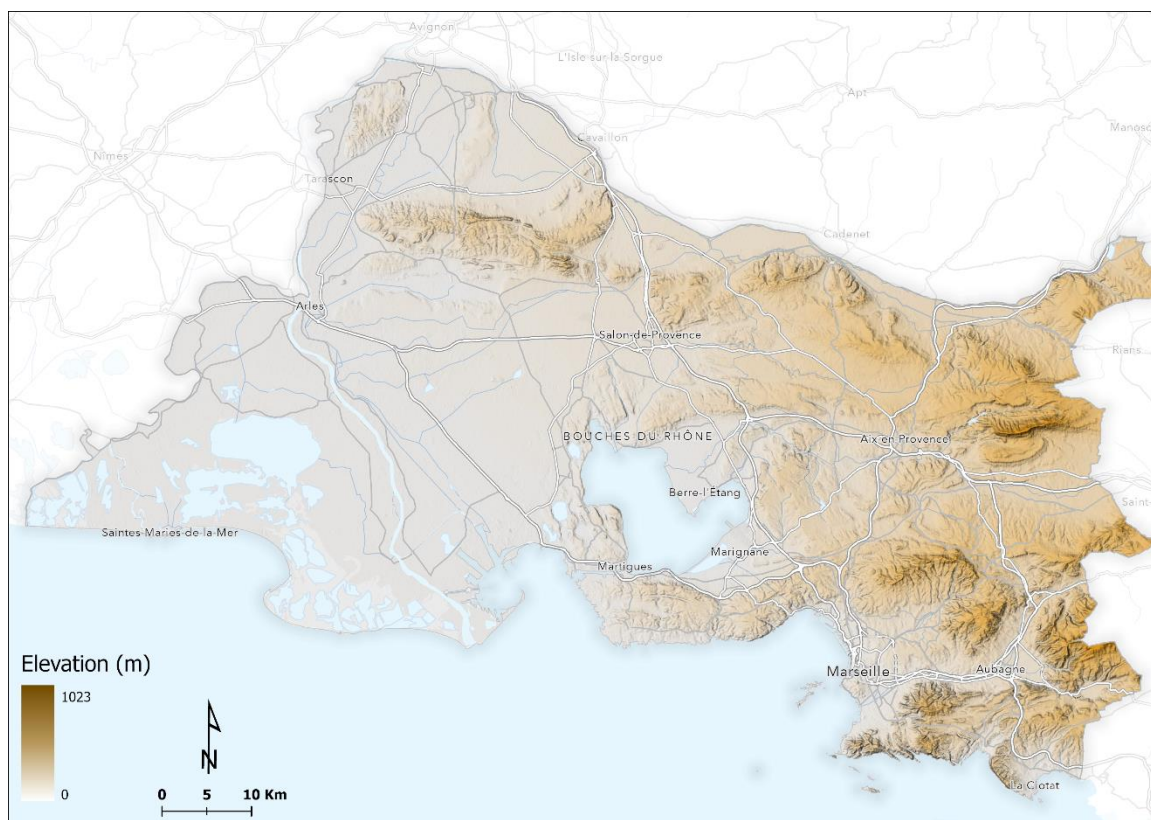


Figure 1 Departmental limits of Bouches-du-Rhône overlaid on a 25 m Digital Elevation Model.

Table 1 Environmental characteristics of Bouches-du-Rhône.

Total area (km ²)	3456
Forested area (km ²)	1530
Mean slope (°)	8.8
Median slope (°)	5.7
Mean elevation (m)	142
Median elevation (m)	89

When considering only fires with a known cause, 63 % of the total burned area and approximately half (51 %) of all fire ignitions in the department are due to arson fires, according to Prométhée (Table 2). In addition, most of the large fires (>100 ha) in the study area are caused by arsonists (Figure 2). Even though negligence (professional & personal) is the second most frequent cause of fire ignition, it does not cause a proportionate volume of burned area.

Table 2 Number of fires and volume of burned area per ignition cause from 1973 to 2020 in Bouches-du-Rhône.

Fire Ignitions (#)	Percentage (%)	Burned Area (ha)	Percentage (%)	Cause
349	3.3	3,020	3.3	Accidental
161	1.5	335	0.4	Natural
1,034	9.7	5,223	5.7	Negligence
1,556	14.7	14,493	15.7	Arson
7,524	70.8	69,105	75	Unknown
10,624	-	92,176	-	Total

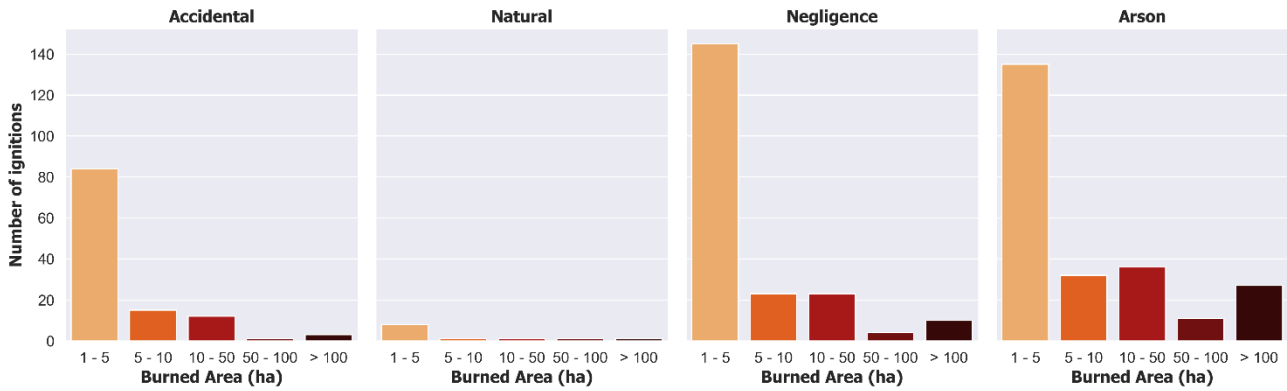


Figure 2 Number of fires per volume category of burned area (> 1ha) and ignition cause from 1973 to 2020.

2.2. Data

2.2.1. Fire database

In France, the majority of forest fire research is based on Prométhée, the national database for forest fires. The specific fire database holds records of fires starting from 1973 and it includes information such as burned area, cause of ignition, date, and approximate location (within a 2x2 km grid) for each fire. In the current study, we used a geographic database that contains exact coordinates of fire ignitions that is provided by the National Forestry Office (Office National des Forêts, ONF), which to the best of our knowledge is the second time being utilized after Ganteaume and Long-Fournel, 2015. The dataset consists of 3,234 fire ignition points ranging from 1960 to 2012, which however does not contain information on the cause of ignition. To enrich the ONF point database with the cause of ignition, two additional databases were used (Figure 3). Information from the Prométhée database was firstly merged with a polygon fire geodatabase that is described in Bountzouklis et al., 2022 and subsequently spatially joined with the point geodatabase resulting in a combined dataset that contains ignition coordinates, burn scars, and cause of ignition.

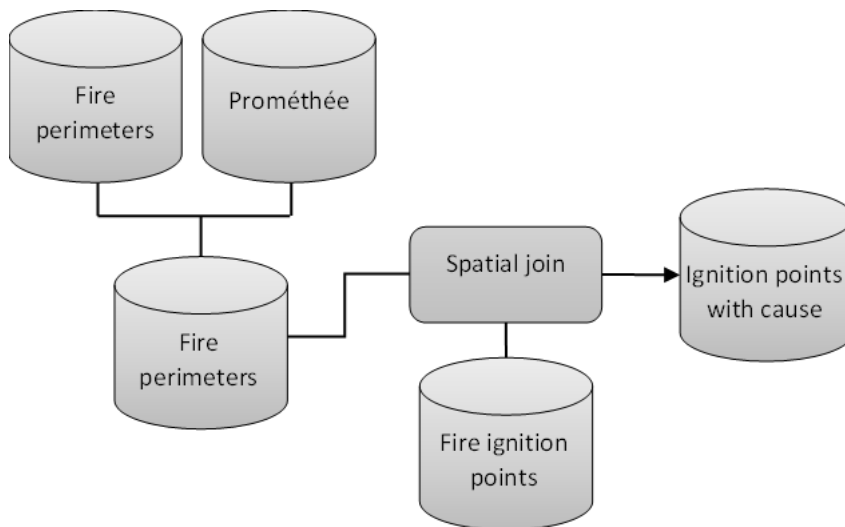


Figure 3 Flow chart depicting the processing steps to generate the final dataset.

As earlier records on fire causes are considered less reliable (Ganteaume and Jappiot, 2013) only fires from 1996 to 2012 were considered, resulting in 323 fires (Figure 4). It was deemed best to classify causes into two major categories, arson and non-arson, due to the limited recorded number of fires caused by accident, negligence and lightning strikes but also due to the significance of arson fires in the specific area.

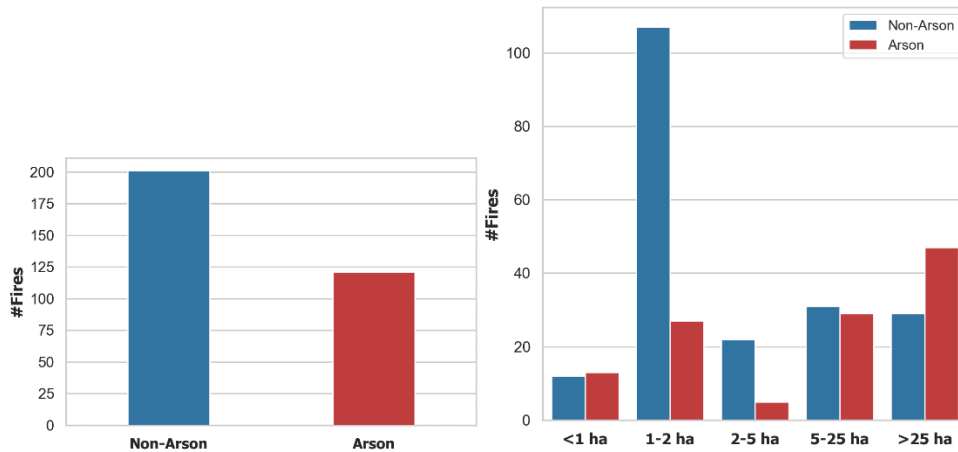


Figure 4 Number of fires per cause and burned area size.

2.2.2. Explanatory variables

Multiple environmental and socioeconomic factors (Table 3 & Figure 5) that are known to be associated with forest fires, were acquired by a combination of European and national databases, in order to train a model that can identify the ignition cause of a fire. To account for any potential geometric errors of the ignition points and more importantly to include contextual geographic information, a circular buffer zone (500 m) was created around each fire ignition point to extract relative information.

Table 3 List of environmental, anthropogenic and spatiotemporal variables considered.

Type	Name	Description	Source
Land Cover	Artificial surfaces	Percentage of cover within each zone	Corine Land Cover - 2006
	Agriculture		
	Vegetation		
	Water/ Wetlands		
Topographic	Slope inclination	Mean inclination (°)	National Geographic Institute - 5 m spatial resolution
	Slope aspect	Percentage of cover within each zone (Flat, N, E, S, W)	
	Elevation	Mean elevation (m) above sea level	
	Topographic wetness index	Topographic-driven control on soil moisture	
Vegetation type	Coniferous	Percentage of cover within each zone	National Geographic Institute - 2006
	Coppice		
	Hardwood		
	Open Forest		
	Shrublands		
Anthropogenic	Population density	Individuals per sq. m.	National Institute of Statistics and Economic Studies - Sub-municipal level - 2006
	Gini index	Inequality index	
	Income	Mean taxable income (€)	
	Unemployment	Unemployment rate (%)	
	Primary road distance	Euclidean distance from fire location (m)	National Geographic Institute - 2008
	Secondary road distance		
	Power pylons distance		
	Railway distance		
Wildland Urban Interface distance			
House density	Buildings per sq. m.		
Spatiotemporal characteristics	Season	Winter, spring, summer, autumn	Fire database
	Burned area size	<1 ha, 1-2 ha, 1-2 ha, 2-5 ha, 5-25 ha, >25 ha	
	Coordinates	XY coordinates of each fire	

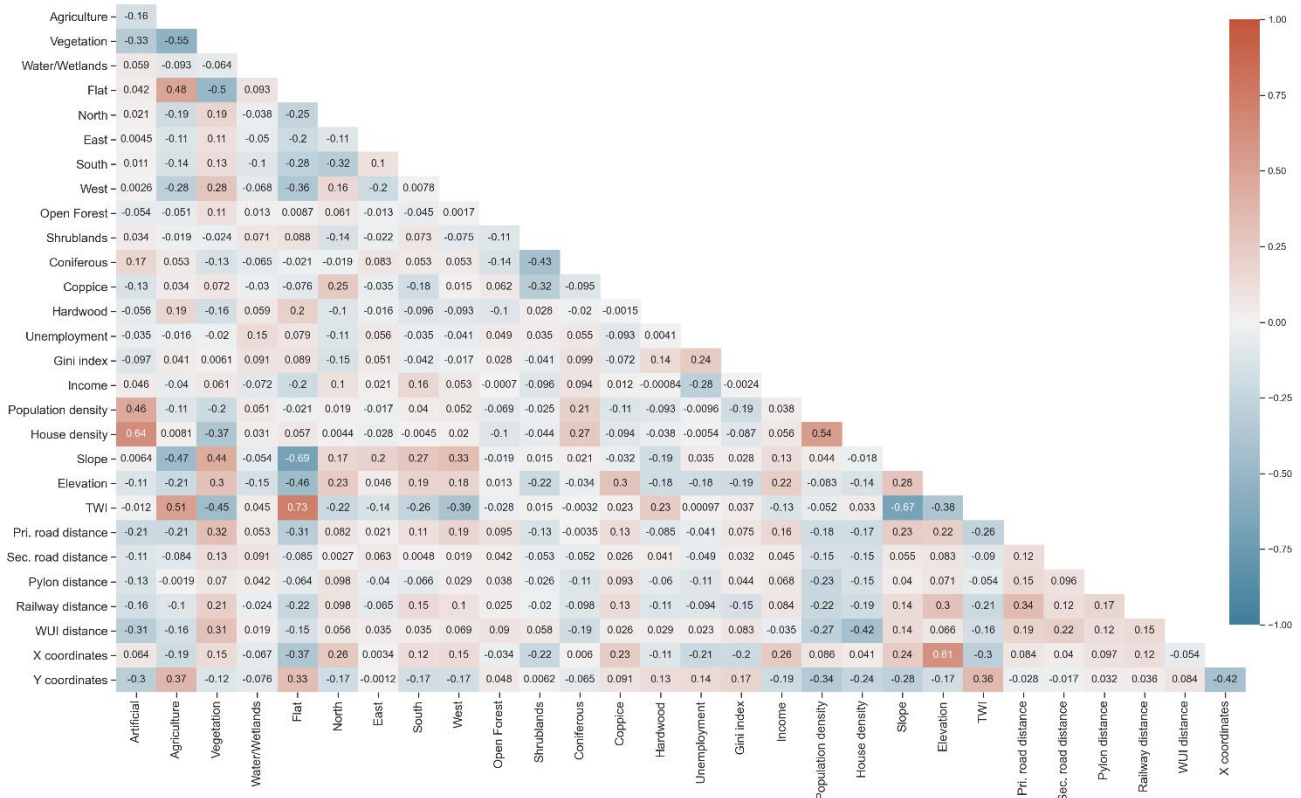


Figure 5 Kendall's Tau rank correlation coefficient heatmap of explanatory variables.

2.3. Methods

The analysis of the data is based on Random Forests (Breiman, 2001), a well-established machine learning algorithm in many disciplines but also in wildfire science (Jain et al., 2020). In order to train the model, 70% of the original dataset was utilized, while the remaining 30% was used for testing the accuracy in predicting the cause of a fire. Due to the limited number of observations, the process of splitting the data (using the same ratio) and executing the model was iterated (n=300) to have a more consistent perception of the accuracy of the model. Additionally, the processing chain included tuning the algorithm's hyperparameters as well as calculating the feature importance score for all variables. Feature importance was based on the Gini impurity method which can be used as a diagnostic which contributes to understanding which variables are driving the results of a model and which ones can be discarded.

3. Results

The accuracy of the model is illustrated in the form of a boxplot (Figure 6) that represents accuracy scores from all iterations of the model. The accuracy of the model to classify unknown caused fires can vary substantially ranging from 56% to 76% (median value 67%), due to the small size of the dataset.

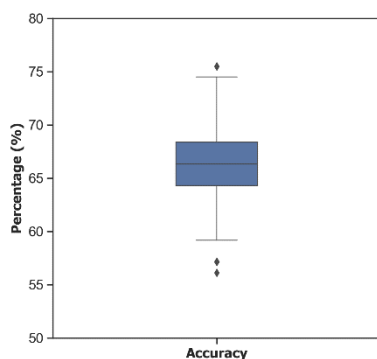


Figure 6 Boxplot representing Random Forest's classification accuracy for all iterations (n=300). (i) Bar within the box is the median value, (ii) bottom part of the box is the first quartile, and (iii) top part of the box is the third quartile. Whiskers represent observations outside the middle 50% and points represent outliers.

Figure 7 shows the importance values of each explanatory variable used in the model. Overall, anthropogenic features (in blue) appear to surpass in importance the rest. Moreover, topographic factors (in brown) seem to be more important than land cover and vegetation type while spatiotemporal variables except for XY coordinates hold the lowest importance. More specifically, Secondary road distance, Shrublands and Unemployment rate are the three variables displaying the highest importance. Finally, several variables mainly related to spatiotemporal characteristics and others such as certain vegetation types (Hardwood & Coppice) as well as Water can be excluded to decrease model complexity and execution time.

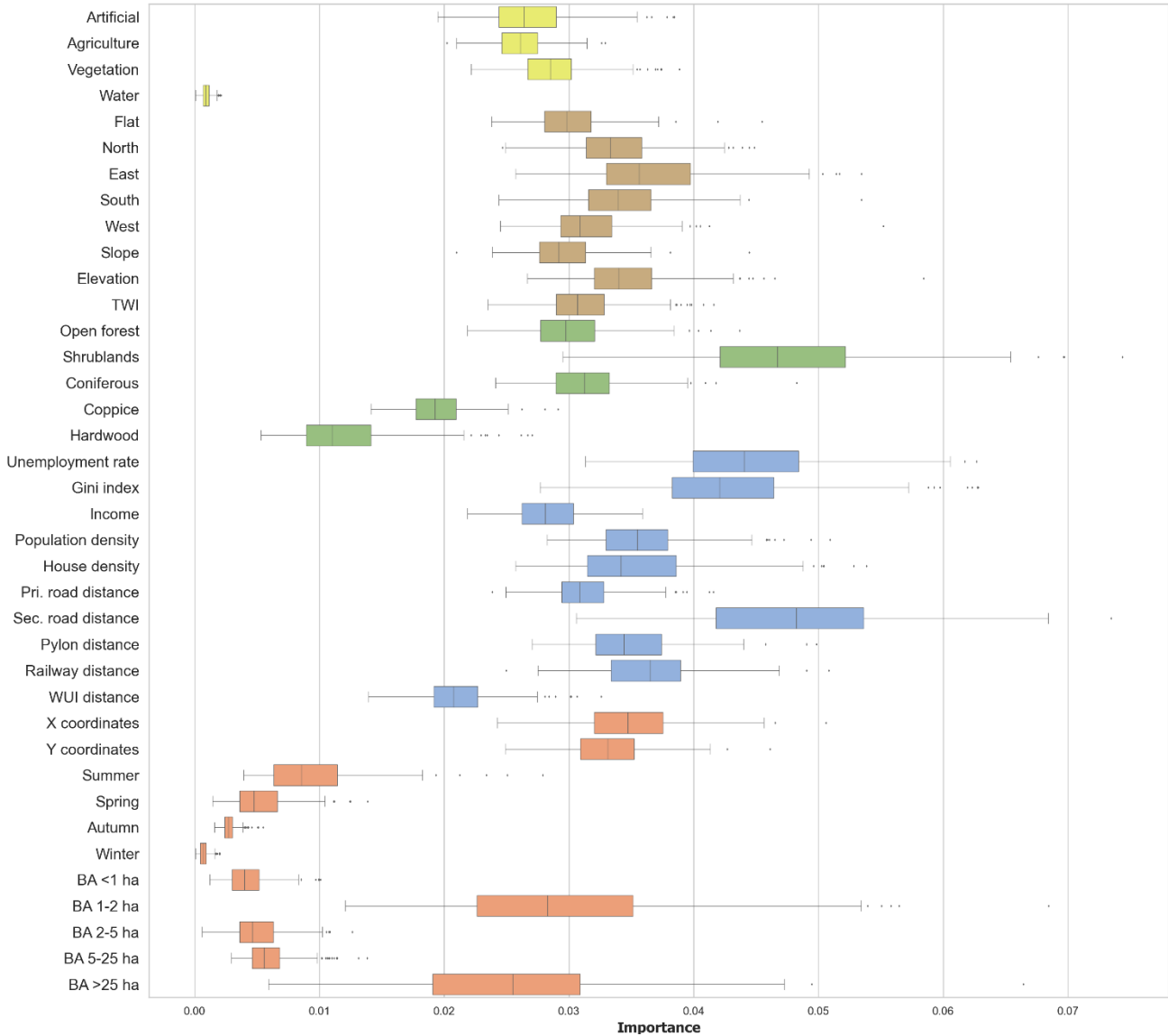


Figure 7 Distribution of the variable importance values for all iterations (n=300).

4. Conclusion

The results of the study suggest that the source of unknown caused fires can be identified at an acceptable level of accuracy even with a limited number of fires. Anthropogenic drivers such as distance to secondary roads and unemployment rate, along with higher volumes of shrublands around ignition points are the most important features in determining the classification of unknown caused fires for the specific area.

Overall performance of such models would most likely greatly benefit from the exploitation of larger datasets as well as from the inclusion of weather-related variables. Finally, as location holds particular importance over certain fire causes, spatial extensions of machine learning algorithms such as Geographic random forests

(Georganos et al., 2021) and Geographically weighted neural networks (Hagenauer and Helbich, 2022) could provide significant enhancements over the original algorithms.

5. References

- Bountzouklis, C., Fox, D. M. and Di Bernardino, E.: Environmental factors affecting wildfire-burned areas in southeastern France, 1970–2019, *Nat. Hazards Earth Syst. Sci.*, 22(4), 1181–1200, doi:10.5194/nhess-22-1181-2022, 2022.
- Breiman, L.: Random Forests, *Mach. Learn.*, 45, 5–32, doi:https://doi.org/10.1023/A:1010933404324, 2001.
- Catry, F. X., Rego, F. C., Bação, F. L. and Moreira, F.: Modeling and mapping wildfire ignition risk in Portugal, *Int. J. Wildl. Fire*, 18(8), 921–931, doi:10.1071/WF07123, 2009.
- Curt, T., Fréjaville, T. and Lahaye, S.: Modelling the spatial patterns of ignition causes and fire regime features in southern France: Implications for fire prevention policy, *Int. J. Wildl. Fire*, 25(7), 785–796, doi:10.1071/WF15205, 2016.
- Ganteaume, A. and Jappiot, M.: What causes large fires in Southern France, *For. Ecol. Manage.*, 294, 76–85, doi:10.1016/j.foreco.2012.06.055, 2013.
- Ganteaume, A. and Long-Fournel, M.: Driving factors of fire density can spatially vary at the local scale in south-eastern France, *Int. J. Wildl. Fire*, 24(5), doi:10.1071/WF13209, 2015.
- Georganos, S., Grippa, T., Niang Gadiaga, A., Linard, C., Lennert, M., Vanhuysse, S., Mboga, N., Wolff, E. and Kalogirou, S.: Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling, *Geocarto Int.*, 36(2), 121–136, doi:10.1080/10106049.2019.1595177, 2021.
- Gonzalez-Olabarria, J. R., Brotons, L., Gritten, D., Tudela, A. and Teres, J. A.: Identifying location and causality of fire ignition hotspots in a Mediterranean region, *Int. J. Wildl. Fire*, 21(7), 905, doi:10.1071/WF11039, 2012.
- Hagenauer, J. and Helbich, M.: A geographically weighted artificial neural network, *Int. J. Geogr. Inf. Sci.*, 36(2), 215–235, doi:10.1080/13658816.2021.1871618, 2022.
- Jain, P., Coogan, S. C. P., Subramanian, S. G., Crowley, M., Taylor, S. and Flannigan, M. D.: A review of machine learning applications in wildfire science and management, , doi:10.1139/er-2020-0019, 2020.
- Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A. and Pereira, J. M. C.: Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest, *For. Ecol. Manage.*, 275, 117–129, doi:10.1016/j.foreco.2012.03.003, 2012.
- Penman, T. D., Bradstock, R. A. and Price, O.: Modelling the determinants of ignition in the Sydney Basin, Australia: Implications for future management, *Int. J. Wildl. Fire*, 22(4), 469–478, doi:10.1071/WF12027, 2013.
- Rodrigues, M., de la Riva, J. and Fotheringham, S.: Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression, *Appl. Geogr.*, 48, 52–63, doi:10.1016/j.apgeog.2014.01.011, 2014.
- Syphard, A. D. and Keeley, J. E.: Location, timing and extent of wildfire vary by cause of ignition, *Int. J. Wildl. Fire*, 24(1), 37–47, doi:10.1071/WF14024, 2015.
- Syphard, A. D., Radeloff, V. C., Keuler, N. S., Taylor, R. S., Hawbaker, T. J., Stewart, S. I. and Clayton, M. K.: Predicting spatial patterns of fire on a southern California landscape, *Int. J. Wildl. Fire*, 17(5), 602–613, doi:10.1071/WF07087, 2008.

LETTER • OPEN ACCESS

Predicting wildfire ignition causes in Southern France using eXplainable Artificial Intelligence (XAI) methods

To cite this article: Christos Bountzouklis *et al* 2023 *Environ. Res. Lett.* **18** 044038

View the [article online](#) for updates and enhancements.

You may also like


- [Analysis of Weed Growth in Rabi Crop Agriculture Using Deep Convolutional Neural Networks](#)
Anand Muni Mishra, Prabhjot kaur, Yogesh Shahare et al.
- [ExoMiner: A Highly Accurate and Explainable Deep Learning Classifier That Validates 301 New Exoplanets](#)
Hamed Valizadegan, Miguel J. S. Martinho, Laurent S. Wilkens et al.
- [Depression evaluation based on prefrontal EEG signals in resting state using fuzzy measure entropy](#)
Feifei Chen, Lulu Zhao, Baimin Li et al.

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Predicting wildfire ignition causes in Southern France using
eXplainable Artificial Intelligence (XAI) methods

OPEN ACCESS

RECEIVED
8 February 2023REVISED
28 March 2023ACCEPTED FOR PUBLICATION
30 March 2023PUBLISHED
11 April 2023Christos Bountzouklis^{1,*} , Dennis M Fox¹ and Elena Di Bernardino²¹ UMR CNRS 7300 ESPACE, Université Côte d'Azur, Nice, France² UMR CNRS 7351 LJAD, Université Côte d'Azur, Nice, France

* Author to whom any correspondence should be addressed.

E-mail: christos.bountzouklis@univ-cotedazur.fr**Keywords:** wildfires, machine-learning, eXplainable Artificial Intelligence, ignition causes, Southern FranceOriginal content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.**Abstract**

The percentage of wildfires that are ignited by an undetermined origin is substantial in Europe and Mediterranean France. Forest fire experts have recognized the significance of fires with an unknown ignition source since documentation and research of fire causes are important for creating appropriate fire policies and prevention strategies. The use of machine learning in wildfire science has increased considerably and is driven by the increasing availability of large and high-quality datasets. However, the absence of comprehensive fire-cause data hinders the utility of existing fire databases. This study trains and applies a machine-learning based model to classify the cause of fire ignition based on several environmental and anthropogenic features in Southern France using an eXplainable Artificial Intelligence framework. The results demonstrate that the source of unknown caused wildfires can be predicted at various levels of accuracy/natural fires have the highest accuracy (F1-score 0.87) compared to human-caused fires such as accidental (F1-score 0.74) and arson (F1-score 0.64). Factors related to spatiotemporal properties as well as topographic characteristics are considered the most important features in determining the classification of unknown caused fires for the specific area.

1. Introduction

In Europe, approximately 50% of all fires were caused by an unknown origin based on data reported from 19 European countries in the European Forest Fire Database from 1999 to 2016 (De Rigo *et al* 2017). In Mediterranean France, almost 70% of all fires between 1973 and 2020 were recorded without a cause of ignition according to the forest fire database for the Mediterranean area (Prométhée.com). Many experts in the field of fire management in Europe have acknowledged the importance of fires classified as having an undetermined origin (Tedim *et al* 2022), since the lack of information regarding fire causes makes it difficult for fire managers to determine the most suitable course of action to prevent similar incidents from happening in the future. French fire experts, in particular, have identified fires of unknown origin as being of paramount importance (Tedim *et al*

2022) among the various categories of fires: natural, accident, negligence, deliberate, and rekindle of the harmonized classification scheme of fire causes in Europe (Camia *et al* 2013). In southeastern France, Ganteaume and Guerra (2018) highlighted the fact that large areas are burned by fires of undetermined sources, and they argue for enhanced quality and quantity of investigations into fire ignition causes in order to improve the accuracy of fire databases. Fire ignition patterns can vary significantly both temporally and spatially depending on the cause of ignition (Curt *et al* 2016) and can be impacted by a plethora of environmental and anthropogenic drivers (Syphard *et al* 2008, Catry *et al* 2009, Syphard and Keeley 2015). As such, documentation and research of fire causes and their spatiotemporal patterns are essential for establishing meaningful fire policies (Rodrigues *et al* 2014) since a better understanding of these patterns can improve the efficacy of fire prevention

strategies (Oliveira *et al* 2012). However, the absence of comprehensive fire-cause data hampers the utility of these databases.

Similar to other areas of study, the utilization of machine learning (ML) methodologies in wild-fire science has seen a marked increase in popularity in recent years (Jain *et al* 2020, Bot and Borges 2022). Applications of ML in wildfire science include classifying and mapping fuel properties (Riley *et al* 2014) and quantities (López-Serrano *et al* 2016) as input variables in fire behavior models; fire and smoke detection (Zhao *et al* 2018, Ba *et al* 2019) for rapid firefighting response; active or final burned area and severity mapping based on remote sensing data for suppression planning and damage assessment (Pereira *et al* 2017, Collins *et al* 2018, Ban *et al* 2020); forecasting of fire weather variables (Bates *et al* 2017, Lagerquist *et al* 2017); prediction of fire occurrence (Dutta *et al* 2016, Vecín-Arias *et al* 2016), fire spread/growth rates (Chetehouna *et al* 2015) and burned area (Mayr *et al* 2018, Hodges and Lattimer 2019) for more effective dissemination of suppression assets; modeling and mapping of fire risk based on various environmental and anthropogenic variables, to identify potential drivers of fire ignition and/or burned area (Curt *et al* 2016, Ghorbanzadeh *et al* 2019, Molina *et al* 2019).

While ML models have demonstrated great effectiveness at identifying complex patterns in large datasets, some are considered ‘black boxes’ because it can be difficult to understand how the model arrives at its predictions or how certain patterns were identified (Loyola-Gonzalez 2019). This lack of interpretability can be a barrier to adoption, as it may be difficult for stakeholders to trust such models without understanding the complete algorithm inference pattern. In recent years, eXplainable Artificial Intelligence (XAI)/interpretable machine learning has emerged as an approach that employs various techniques and strategies to enhance the interpretability, transparency, and explainability of ML models and their decision-making processes with the ultimate goal of fostering trust and accountability in the model’s output. In the context of wildfire science, the application of XAI has been explored by only two recent studies to address wildfire occurrence and size (Al-Bashiti and Naser 2022, Cilli *et al* 2022).

Research conducted on fire ignition causes is fairly limited and poorly understood, but some studies have demonstrated that arson fires can potentially be predicted both spatially and temporally (Gonzalez-Olabarria *et al* 2012, Penman *et al* 2013). The objective of this study is to develop a ML-based model that can classify the ignition source of fires that have been recorded without a known cause in France. Furthermore, this study aims to evaluate the significance and the effect of various environmental and anthropogenic factors in determining the classification of different fire sources utilizing XAI methods.

2. Method

2.1. Study area

The study area comprises 15 administrative divisions (departments) in the south of France, with a total area of just over 80 000 km² (table 1, figure 1). The specific region is considered the most fire-prone in France and where most of the burned area is recorded, despite exhibiting decreasing trends in the last decades (Bountzouklis *et al* 2022). Environmental characteristics and landscapes vary significantly with both mountainous and coastal zones contained in the study area; the highest altitudes and steepest slopes are found in the northeastern parts where the French Alps are located (e.g. Hautes-Alpes, Alpes-de-Haute-Provence) whereas in the southern portions topography is low-lying and relatively flat (e.g. Bouches-du-Rhône, Hérault). Population density is influenced by topography: the highest concentrations are located in areas with low altitudes and gentle slopes, especially in the southeastern Mediterranean coastal and near coastal zones (e.g. Bouches-du-Rhône, Alpes-Maritimes). The French Alps and the island of Corsica are largely covered by forests & semi-natural areas whereas the largest agricultural areas are concentrated mainly in the center of the study area.

2.2. Fire database

The current study was based on ‘Prométhée’, the official forest fire database for Mediterranean area in France. This database documents fires from 1973 onwards and contains information for each fire such as burned area, ignition source (known/unknown), time, date, and location within a 2 × 2 km grid. Similar to the harmonized European classification scheme on ignition causes (Camia *et al* 2013), ‘Prométhée’ includes five major fire ignition sources: (i) accidental (e.g. power lines, vehicles), (ii) arson (e.g. pyromania, conflict), (iii) private negligence (e.g. cigarette butts, leisure), (iv) professional negligence (e.g. industry, agriculture) and (v) lightning. The total number of fires considered in our study is 48 038; these were recorded from 1997 to 2020. Fire records prior to 1997 were excluded from this study since classification on the origin of fires is considered less reliable (Ganteaume and Jappiot 2013). The dataset comprised of records starting in 1997 is fairly balanced with regards to the number of fires of known/unknown sources as approximately 60% have a known cause of ignition. Within the known causes ($n = 27\,620$) frequency varies considerably; arson is the most frequent (38.4%), followed by private negligence (26.7%), professional negligence (17.2%), accidental (10.1%) and finally lightning (7.6%) (figure 2).

2.2.1. Fire frequency & burned area according to cause

After unknown causes, arson fires are both the most numerous and account for the greatest annual burned

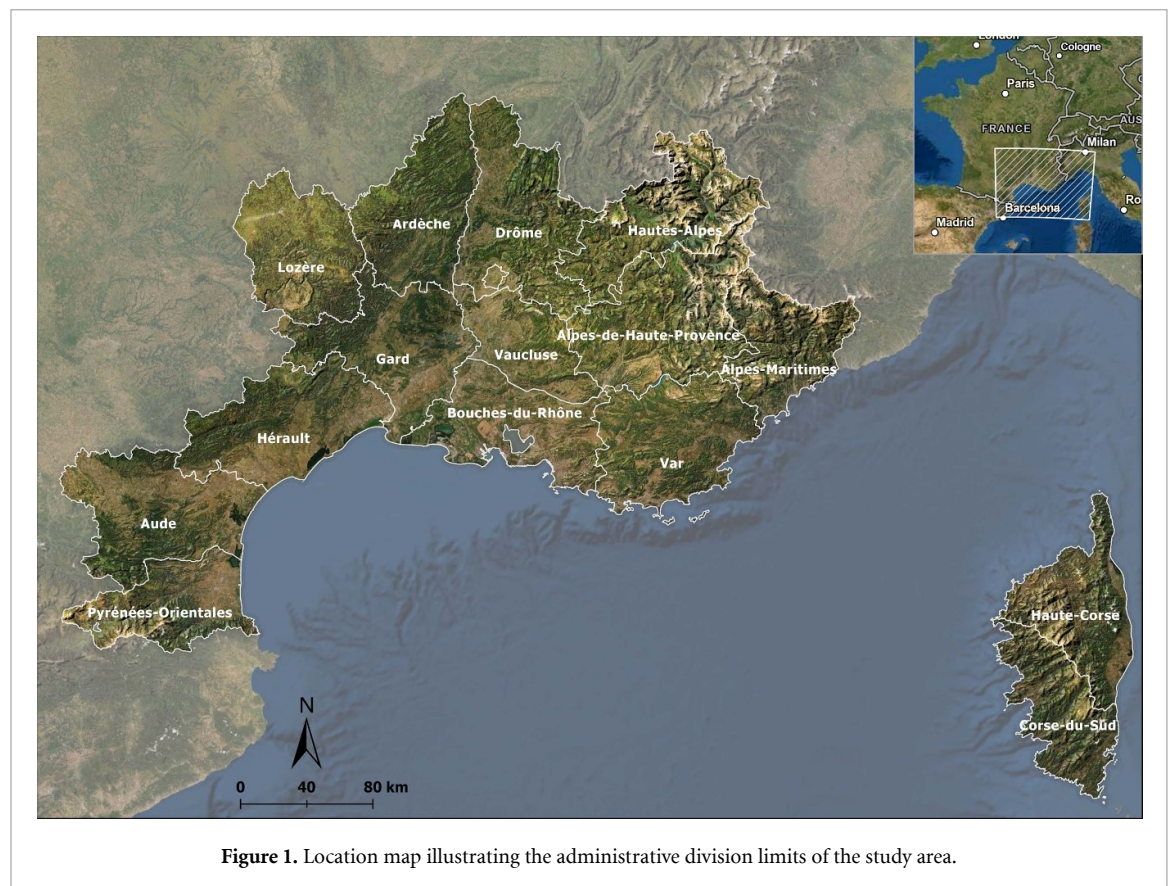
Table 1. Physical and human characteristics of the administrative divisions within the study area based on data from.

Administrative division	Area ^a (km ²)	Mean elevation ^a (m)	Mean slope ^a (°)	Population density ^c (people/km ²)	Artificial surfaces ^b (%)	Agricultural areas ^b (%)	Forests & semi-natural areas ^b (%)
Alpes-de-Haute-Provence	6995	1150	15.5	23.7	1.1	19.2	79.7
Hautes-Alpes	5691	1665	20.6	24.8	1.3	15.2	83.5
Alpes-Maritimes	4294	1110	20.3	257.0	8.9	4.5	86.6
Corse-du-Sud	4019	536	15.3	40.4	2.2	11.1	86.7
Haute-Corse	4707	592	16.2	39.8	1.9	12.4	85.6
Ardèche	5566	596	11.1	59.4	2.7	28.1	69.3
Aude	6344	359	7.4	59.9	2.8	48.4	48.7
Bouches-du-Rhône	5091	142	3.7	405.1	15.7	42.2	42.1
Drôme	6559	581	11.1	80.0	3.4	41.1	55.5
Gard	5875	248	6.2	128.2	6.5	40.3	53.2
Hérault	6230	264	6.5	195.5	7.2	41.5	51.3
Lozère	5176	1025	9.3	14.8	0.7	25.5	73.8
Pyrénées-Orientales	4139	855	12.6	117.3	5.2	27.5	67.3
Var	6032	364	7.8	181.6	9.2	20.9	69.9
Vaucluse	3578	339	5.6	157.3	6.8	53.8	39.4

^a French National Institute of Geographic (IGN).

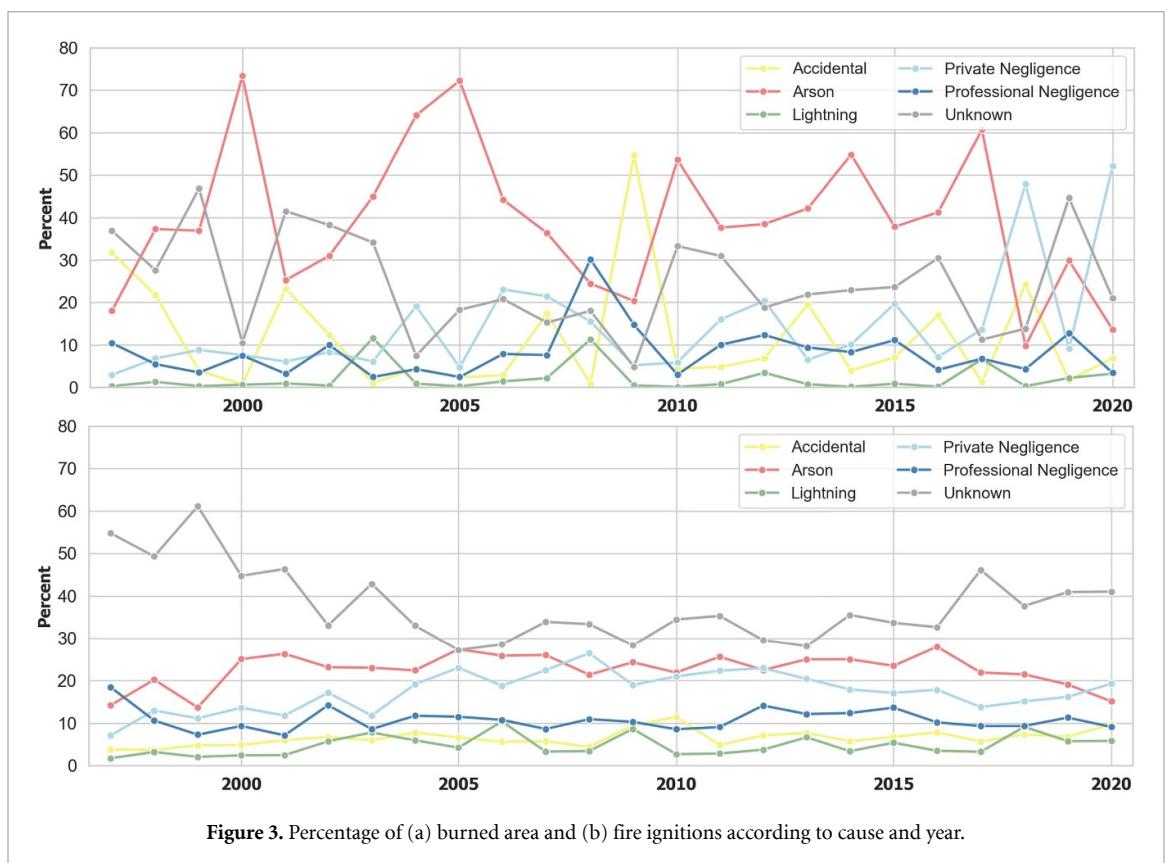
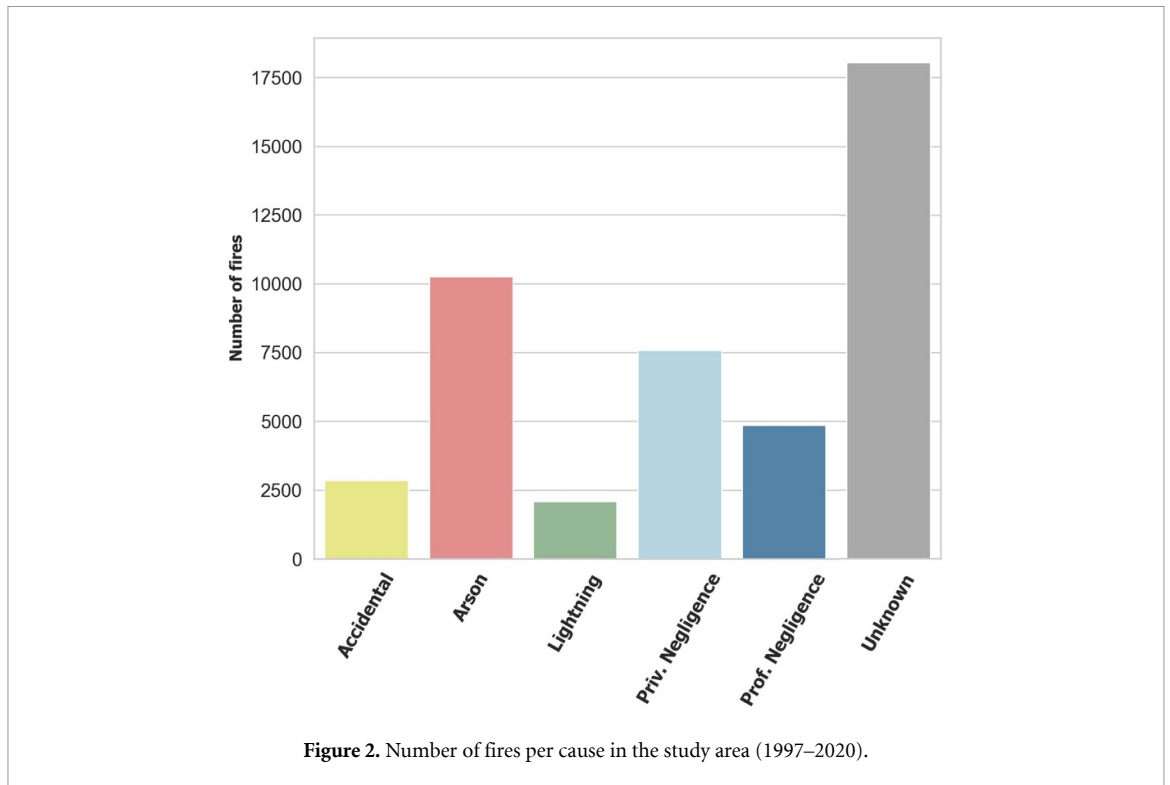
^b Corine Land Cover (2018).

^c National Institute of Statistics and Economic Studies (INSEE).

**Figure 1.** Location map illustrating the administrative division limits of the study area.

area most years (figures 3(a) and (b)). This is followed by private negligence, which, even though is the second most frequent fire source, it does not cause a proportionate extent of burned area. Despite similar numbers of accidental and lightning fires, the annual percentage of area burned by accidental fires is often

significantly greater than that burned by lightning fires and occasionally greater than the other causes. Lastly, although the percentage of burned areas by unknown origin fires is substantial most years, frequently second after arson, it fluctuates widely from 5% to 49% depending on the year.



2.3. Considered features

Table 2 describes the land cover, topographic, anthropogenic, and spatiotemporal variables that were used as features to predict the fire ignition source. The contextual geographic information of the selected

factors was processed for each 2 × 2 km grid initially in ArcGIS Pro v2.9 and subsequently using python packages pandas (McKinney 2010) and NumPy (Harris et al 2020) to preprocess the data for the classification scheme (e.g. replace missing values, one-hot

Table 2. List of features used to model the occurrence of forest fires.

Type	Name	Description	Source
Land cover	Artificial surfaces	Percentage of cover within each grid	Corine Land Cover—2006 (Raster—100 m spatial resolution)
	Agriculture		
	Forest		
	Low vegetation (e.g. shrublands)		
	Water		
Topographic	Wetlands	Mean elevation (m) above sea level	National Geographic Institute (Raster—100 m spatial resolution)
	Elevation		
	Slope aspect		
	Topographic wetness index (TWI)		
Anthropogenic	Topographic ruggedness index (TRI)	Topographic-driven control on soil moisture	National Institute of Statistics and Economic Studies—Sub-municipal level—2006—(Tabular)
	Population density	Number of individuals/area	
	Gini index	Inequality index	
	Income	Mean taxable income (€)	
	Unemployment	Unemployment rate (%)	
	Primary road density	Total line length/area	
	Secondary road density		
Powerline density			
Spatiotemporal	Railway density		National Geographic Institute—2008 (Vector)
	Season	Winter (Dec–Feb), spring (Mar–May), summer (Jun–Aug), autumn (Sep–Nov)	
	Burned area size	<0.1 ha, 0.1–0.5 ha, 0.6–1 ha, 2–5 ha, 1–25 ha, >25 ha	
	Coordinates	XY centroid coordinates	
	Time of the day	Morning (5:00–12:00), afternoon (12:00–17:00), evening (17:00–21:00), night (21:00–5:00)	Fire database ‘Prométhée’

encoding, etc) and finally for visualization purposes seaborn (Waskom 2021).

2.4. Fire cause classification based on random forests (RFs)

ML methodologies learn and adapt through the process of experience, where the size and quality of the input data play a critical role in determining the overall effectiveness of the model. RFs (see e.g. Breiman 2001) is a supervised ML algorithm used both for classification and regression that is well-established in many disciplines and has grown substantially in popularity in the field of wildfire science over the last decade (Jain *et al* 2020). RF is based on decision trees (Breiman *et al* 2017), where each decision tree is a series of If-Then-Else sequences with several branches connected by decision nodes and finally by leaf nodes

that eventually determine a value or category such as the label of a classification task (figure 4). Furthermore, a fundamental characteristic of RF is that a random subset of features is used at each node of each decision tree, resulting in several individually trained and uncorrelated decision trees, and these are finally merged into a larger ensemble model to limit overfitting and produce more accurate predictions.

The processing chain of RF (classification, accuracy score, confusion matrix, hyperparameter tuning, etc) was carried out using the implementation of the algorithm in Python module Scikit-Learn (Pedregosa *et al* 2011) (figure 5). To address the unbalanced number of samples between classes, the Synthetic Minority Oversampling Technique (SMOTE) (Chawla *et al* 2002) was used, which is implemented under Python package scikit-learn

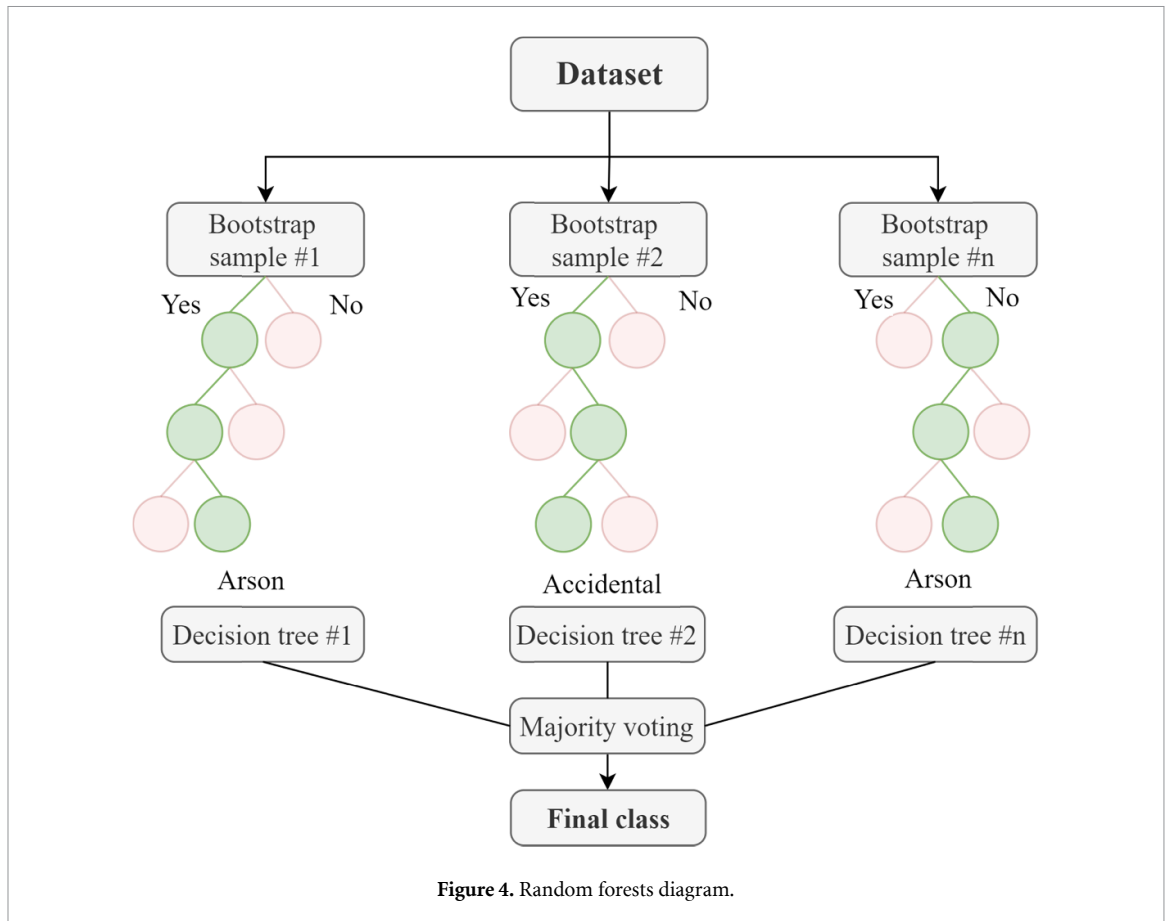


Figure 4. Random forests diagram.

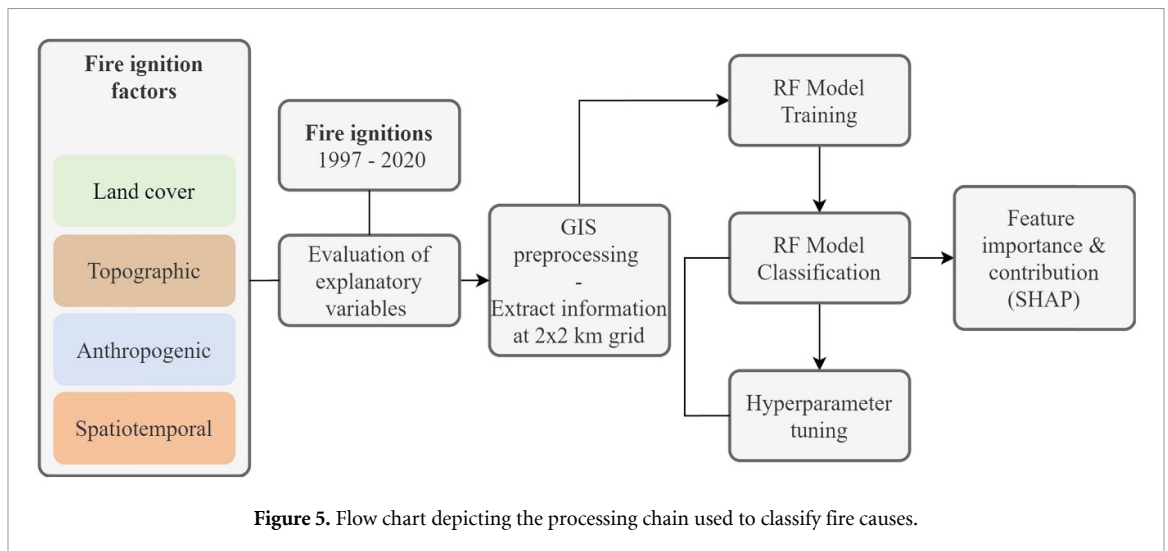


Figure 5. Flow chart depicting the processing chain used to classify fire causes.

imbalanced-learn (Lemaitre *et al* 2017). SMOTE is a common method to produce synthetic data from a minority class (e.g. lightning ignitions) by randomly selecting one of the k-nearest-neighbors and using it to generate new, but randomly tweaked, similar samples. To train the classifier, 70% of the dataset was utilized, while the remaining 30% was used for testing the accuracy in predicting the cause of a fire. The synthetic samples created using SMOTE were utilized only during the training phase and not for

the validation of the model. To finetune the algorithm hyperparameters such as number of trees, max number of features considered for splitting a node, max levels in each decision tree etc, scikit-learn Random Search Cross Validation method was used; this allowed us to evaluate and narrow down a wide range of values for each hyperparameter. Subsequently, the Grid Search with Cross Validation method was used to examine different combinations of specific values for the selected hyperparameters.

To evaluate the model, accuracy, precision, recall and F1-score were used; these are commonly used as evaluation metrics for classifiers in the field of ML, which are calculated using the number of instances classified as true positives (TP), true negatives (TN), false positives (FP) or false negatives (FN). Accuracy is determined by dividing the number of correct predictions by the total number of instances. Precision specifies how many of the instances the classifier predicted as positive are actually positive while recall shows what fraction of the positive instances in the dataset were correctly identified by the classifier. F1-score serves as a comprehensive measure of both precision and recall. A F1-score of 1 denotes optimal performance, with both precision and recall being maximized. Conversely, a score of 0 represents the worst possible outcome, with both precision and recall being minimal. A score of 0.5, which is equivalent to random guessing, is suboptimal performance, whereas scores above 0.5 are generally considered to indicate good performance,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total number of instances}} \quad (1)$$

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

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

$$\text{F1score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

To identify which features are driving the classification but also to comprehend the contribution of each one, the SHapley Additive exPlanations (SHAPs) (Lundberg and Lee 2017) method was utilized. SHAP is an approach based on game theory that is used to explain the ML model outputs by breaking down the prediction into contributions from each feature value. These contributions are combined and help us understand the overall importance of each feature value in the final prediction. SHAP values can be visualized using various plots, such as a summary plot, that allow us to display not only the strength of the impact a certain feature has but also the direction of the impact.

3. Results

As elaborated below, the results derived from the RF model are presented through classification metrics and a confusion matrix, subsequently followed by the description of which features drive the classification and how they influence it.

3.1. Fire ignition cause classification

The overall accuracy of the multiclass RF classification scheme reaches about 70% (69.8%). Detailed results per ignition cause are presented in table 3 and evaluated using (i) F1-score, (ii) precision and (iii) recall. Concerning the accidental class, the model displays the second highest F1-score (0.77) and a moderate discrepancy between and precision (0.81) and recall (0.74). This indicates that the model is able to correctly identify most of the instances as accidental when it predicts that class, but it misses more instances that actually belong to that cause. Regarding the arson class, the model shows a lower F1-score of 0.64 and not very accurate in terms of precision (0.60), meaning that it may predict some instances as arson that actually belong to a different class. However, the model performs better when it comes to identifying most of the instances that belong to the arson class (recall score 0.69). The lightning class displays overall the best classification metrics (F1-score of 0.88). The precision score (0.85) is fairly lower than the recall score (0.91), suggesting that lightning fires are easier for the model to identify and are not confused with another class. On the contrary, the model performs the worst for the private negligence class, with a F1-score of 0.55. In this class, the precision score (0.59) is higher than the recall score (0.52), which suggests that the classifier has a higher rate of correctly identifying positive samples but is missing a higher proportion of the total number of positive samples. Finally, the professional negligence class exhibits relatively low but balanced scores between precision (0.67) and recall (0.63).

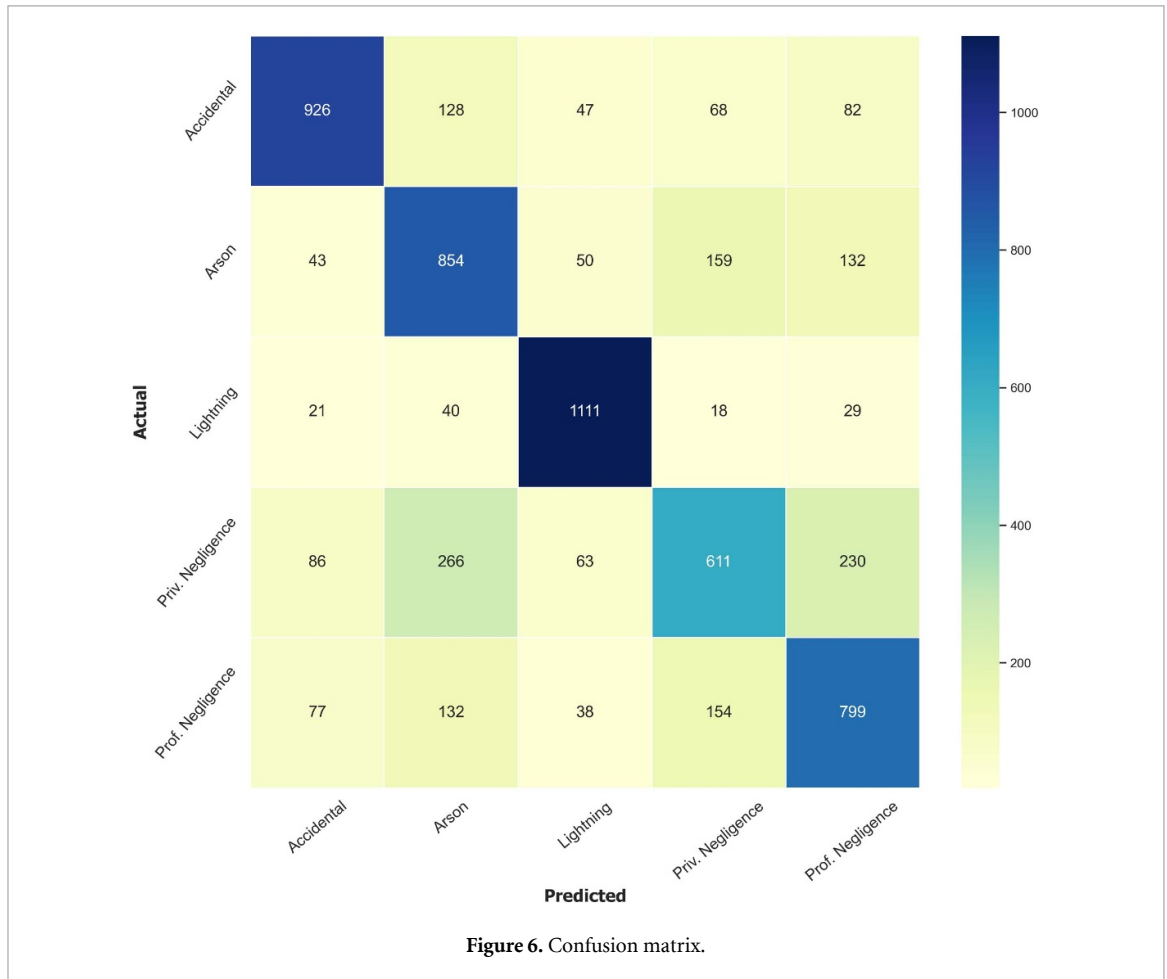
The confusion matrix (figure 6) provides additional information with regards to the performance of the classification of ignition causes. More specifically, accidental fires are most frequently misclassified as arson ones. There is a high number ($n = 159$) of arson fires that are wrongly classified as private negligence, and similarly, there are 266 private negligence fires that are misclassified as arson. This could mean that there are similarities between the causes of these fires, or that the model may not have enough information to accurately distinguish between these classes. As the most accurately predicted cause, lightning displays low misclassification numbers, which are distributed evenly among the other classes. In contrast, private negligence, that is a major negative contributor to the overall classification accuracy, shares its errors primarily between professional negligence and arson classes. Finally, professional negligence fires are also often confused for either arson or private negligence fires.

3.2. Feature importance and effect

Figure 7 illustrates the computed feature importance of the RF model for all classes, which is expressed through mean SHAP values that represent the average

Table 3. Classification metrics per ignition cause.

Ignition cause	Precision	Recall	F1-score	Accuracy
Accidental	0.81	0.74	0.77	69.8%
Arson	0.60	0.69	0.64	
Lightning	0.85	0.91	0.88	
Private negligence	0.59	0.52	0.55	
Professional negligence	0.64	0.68	0.65	



impact of a feature on the model output across all the instances in the dataset. Overall, feature importance values vary significantly both per feature type and ignition cause. The features summer, elevation and afternoon form a group that stands out significantly from the rest, followed by a second cluster with slightly less impactful features such as spring, geographic coordinates, BA <0.1 ha and secondary road density.

In the context of accidental fires, several features demonstrate comparable significance, with afternoon, elevation, summer, and primary road density being slightly more salient than other variables. Similarly, the relevance of features for arson fires is widely distributed, with spatiotemporal characteristics such as summer, night, and location being the most prominent factors. Regarding fires caused by lightning, summer and elevation are by far the most impactful variables followed by secondary road density. In

the case of private negligence, summer exhibits the highest level of importance, although this distinction is not substantially greater than that of other variables, such as afternoon, spring, and secondary road density. Finally, with respect to professional negligence, summer represents the most influential factor by a significant margin, with only burned area (BA) size (<0.1 ha) showing discernible differences from other variables.

Figure 8 depicts the most influential ($n = 10$) features for each class of the model in descending order. Furthermore, the impact of each feature on the ignition cause is also illustrated through the positive or negative SHAP values. These values indicate whether an instance is more or less likely to belong to a particular class depending on the magnitude of the feature values. For example, instances with lower elevation values are more likely to be classified as an accidental or arson ignition, and less likely to be categorized as

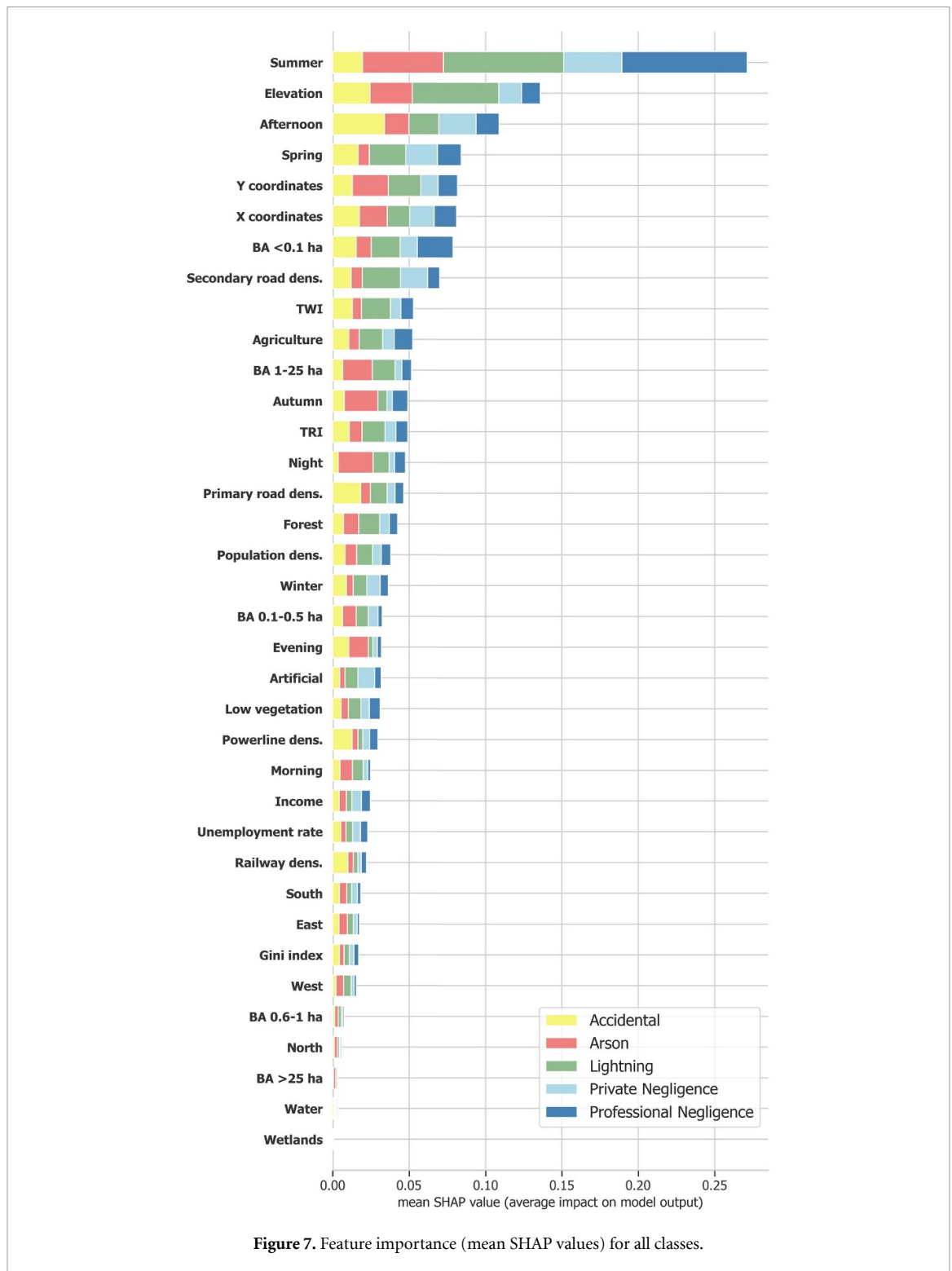


Figure 7. Feature importance (mean SHAP values) for all classes.

a lighting fire. Similarly, a fire that occurred during the summer is more likely to belong to the arson or lightning class, but less probable to be classified into the other categories.

4. Discussion

The performance of the RF classifier varies considerably between natural and human-induced fires. Lightning fires were classified with the highest

accuracy since ignition dynamics for these fires are significantly different from human-caused fires. As reported by Curt *et al* (2016), lightning fires tend to have small burned areas, occur on steep, densely vegetated, mountainous slopes with low anthropogenic presence; seasonality also plays a significant role in the incidence of those fires (summer). This particular profile, which aligns with the interpretation of features effects through the SHAP values, enables the classifier to distinguish it from other causes more clearly.

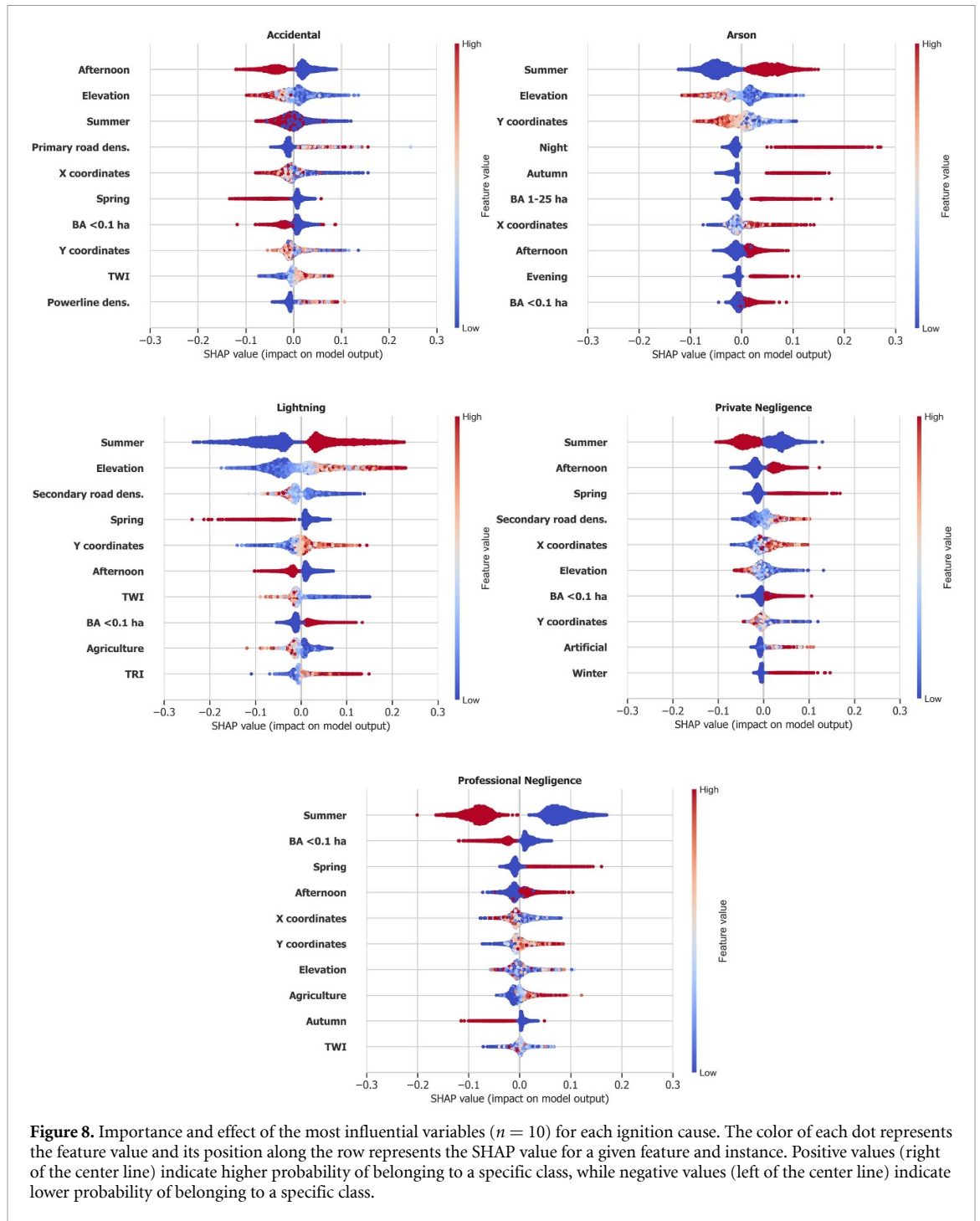


Figure 8. Importance and effect of the most influential variables ($n = 10$) for each ignition cause. The color of each dot represents the feature value and its position along the row represents the SHAP value for a given feature and instance. Positive values (right of the center line) indicate higher probability of belonging to a specific class, while negative values (left of the center line) indicate lower probability of belonging to a specific class.

In contrast to natural fires, human-caused ignitions are multi-faceted and more complex to model. Accidental fires are the least difficult human induced events to classify in our model, potentially attributed to the greater association of such fires with infrastructure elements such as power lines and railways in contrast to other forms of anthropogenic causes. The most challenging cause to classify is private negligence, which is most often misclassified as arson and vice versa. Both arson and private negligence fires often occur in similar contexts, specifically the wildland urban interface. The similarity in environmental contexts and conditions between these types

of fires may make it difficult to distinguish between the two causes. However, this may also reflect a problem of reliability in the fire databases (Ganteaume and Guerra 2018): in order to reduce the number of unknown caused fires the cause is either speculated or attributed without much physical evidence to support it (Camia et al 2013). Professional negligence fires are also confused, but to a lesser extent, with private negligence. Both causes share common characteristics, as they tend to burn small/medium areas and occur mainly outside of the summer season (Curt et al 2016) which is reflected in the significance and impact those features hold in the SHAP framework.

Socioeconomic data used in our model only pertains to a single year. While this approach may have its advantages, such as simplifying data collection and analysis, it can also potentially undervalue the importance of socioeconomic features by not capturing their temporal fluctuations, especially considering that most fires in France but also in the Euro-Mediterranean (95%) region are caused by humans (Ganteaume *et al* 2013, Ganteaume and Jappiot 2013). Factors such as population density, unemployment rate, etc represent dynamic phenomena that can change considerably over time in contrast to static variables such as topography or even to other dynamic variables as, for instance, land cover. The addition of geographic coordinates in our workflow not only partly tackles spatial non-stationarity, as the decision trees of the model in a way incorporate geographic space during their creation, but also enhances the results which is in accordance with other works that utilize ML algorithms for applications of spatial nature (Hengl *et al* 2018). Spatial approaches of ML algorithms such as Geographic Random Forests (Georganos *et al* 2021) and Geographically Weighted Neural Networks (Hagenauer and Helbich 2022) would be advantageous for such applications considering the significance of spatial location and its strong links with different fire ignition causes.

As a first attempt, the current study utilized only the first-level causes (5 categories) available from the hierarchical structure of the 'Promethee' fire database, which also includes second-level (15 categories) and third-level (31 categories) causes. Applying a similar procedure on selected sub level data could possibly improve functionality and understanding of ignition sources and their performance within the classification scheme. However, this would increase the complexity of the model and may negatively impact overall accuracy. Finally, the inclusion of fuel type characteristics and fire-weather variables can potentially strengthen and facilitate the distinction between different fire causes; for instance, arson fires burn larger areas (Ganteaume and Jappiot 2013, Syphard and Keeley 2015) and this may indicate these fires are set under more favorable weather conditions.

The practicality of this model is not intended for operational use or as a substitute method to conventional field investigation methods as it cannot provide physical evidence for the proper deduction of the cause of a forest fire. Instead, it is targeted as a method to analyze large-scale fire databases that contain a moderate percentage of unknown caused fires. The ideal balance would be neither too low, as insufficient data would result in a restricted training dataset, nor too high, as that would render the model less useful. Despite the limitations in identifying causes of unknown ignitions, the results can help to facilitate

targeted prevention efforts (Oliveira *et al* 2012). Moreover, the benefits of harmonized classification systems, such as the one proposed by the European Commission (Camia *et al* 2013) are emphasized. By utilizing such schemes, ML models can significantly benefit from increased harmonized data availability, provided that the data's reliability stays at an adequate level. This allows for the combination of historical national fire databases, leading to the development of larger databases with enhanced modeling potential. Other standardized georeferenced data initiatives in Europe (e.g. the European INSPIRE Directive, the European Geodata Infrastructure, etc) which aim to establish a common framework for the management and sharing of geospatial data across Europe are going in this direction.

ML algorithms have become increasingly popular in fire science (Jain *et al* 2020). These algorithms can help identify complex relationships between various factors that contribute to fire occurrences. However, the success of these algorithms relies heavily on the availability of large, high-quality datasets. As fire science continues to advance, access to larger and more comprehensive datasets is becoming increasingly common. This includes georeferenced explanatory feature data which provides important contextual information that can be used to better understand the underlying causes of fires. As these datasets continue to grow in size and quality, ML algorithms will become even more powerful tools for analyzing fire occurrences and fire causes.

5. Conclusion

In this study we train and apply a model to classify fire ignition causes based on several environmental and anthropogenic features using an XAI framework. The results suggest that the source of unknown caused fires can be identified at various levels of accuracy depending on the nature the forest fire (e.g. F1-score lightning 0.87, accidental 0.74, arson 0.64). Spatiotemporal characteristics including geographic location, season, time of the day but also topographic factors like elevation are the most important features in determining the classification of unknown caused fires for the specific area and fire regime studied here. The role of spatial non-stationarity is highlighted through the importance it holds in our processing framework and should be treated by implementing models that utilize spatial approaches of ML algorithms, which are expected to have increased accuracy over the original ones. The increasing availability of large, high-quality datasets is an important factor driving the growth of ML algorithms in wild-fire science and will likely play a critical role in advancing our understanding of fire causes in the coming years.

Data availability statements

The data that support the findings of this study are openly available at the following URL/DOI: [10.6084/m9.figshare.22028015](https://doi.org/10.6084/m9.figshare.22028015) (Bountzouklis 2023).

Acknowledgments

This research has been supported by the French government through the 3IA Côte d'Azur and UCA-JEDI Investments in the Future project managed by the National Research Agency (Grant Nos. ANR-19-P3IA-0002 and ANR-15-IDEX-01). We would like to thank the two anonymous reviewers for their useful comments that helped to improve the quality of the paper.

ORCID iD

Christos Bountzouklis  <https://orcid.org/0000-0002-1837-1796>

References

- Al-Bashiti M K and Naser M Z 2022 Machine learning for wildfire classification: exploring blackbox, eXplainable, symbolic, and SMOTE methods *Nat. Hazards Res.* **2** 154–65
- Ba R, Chen C, Yuan J, Song W and Lo S 2019 SmokeNet: satellite smoke scene detection using convolutional neural network with spatial and channel-wise attention *Remote Sens.* **11** 1702
- Ban Y, Zhang P, Nascetti A, Bevington A R and Wulder M A 2020 Near real-time wildfire progression monitoring with Sentinel-1 SAR time series and deep learning *Sci. Rep.* **10** 1322
- Bates B C, Dowdy A J and Chandler R E 2017 Classification of Australian thunderstorms using multivariate analyses of large-scale atmospheric variables *J. Appl. Meteorol. Clim.* **56** 1921–37
- Bot K and Borges J G 2022 A systematic review of applications of machine learning techniques for wildfire management decision support *Inventions* **7** 15
- Bountzouklis C 2023 Dataset (figshare) (<https://doi.org/10.6084/m9.figshare.22028015.v1>)
- Bountzouklis C, Fox D M and di Bernardino E 2022 Environmental factors affecting wildfire-burned areas in southeastern France, 1970–2019 *Nat. Hazards Earth Syst. Sci.* **22** 1181–200
- Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32
- Breiman L, Friedman J H, Olshen R A and Stone C J 2017 *Classification and Regression Trees* (New York: Routledge) (<https://doi.org/10.1201/9781315139470>)
- Camia A, San-Miguel-Ayanz J and Durrant T 2013 Harmonized classification scheme of fire causes in the EU adopted for the European Fire Database of EFFIS *Publications Office* (<https://doi.org/10.2788/86859>)
- Catry F X, Rego F C, Bação F L and Moreira F 2009 Modeling and mapping wildfire ignition risk in Portugal *Int. J. Wildland Fire* **18** 921–31
- Chawla N V, Bowyer K W, Hall L O and Kegelmeyer W P 2002 SMOTE: synthetic minority over-sampling technique *J. Artif. Intell. Res.* **16** 321–37
- Chetehouna K, Tabach E E, Bouzaoui L and Gascoin N 2015 Predicting the flame characteristics and rate of spread in fires propagating in a bed of Pinus pinaster using artificial neural networks *Proc. Saf. Environ. Prot.* **98** 50–56
- Cilli R et al 2022 Explainable artificial intelligence (XAI) detects wildfire occurrence in the Mediterranean countries of Southern Europe *Sci. Rep.* **12** 16349
- Collins L, Griffioen P, Newell G and Mellor A 2018 The utility of random forests for wildfire severity mapping *Remote Sens. Environ.* **216** 374–84
- Curt T, Fréjaville T and Lahaye S 2016 Modelling the spatial patterns of ignition causes and fire regime features in southern France: implications for fire prevention policy *Int. J. Wildland Fire* **25** 785–96
- de Rigo D, Liberta' G, Houston Durrant T, Artés Vivancos T and San-Miguel-Ayanz J 2017 *Forest fire danger extremes in Europe under climate change: variability and uncertainty Publications Office of the European Union* (<https://doi.org/10.2760/13180>)
- Dutta R, Das A and Aryal J 2016 Big data integration shows Australian bush-fire frequency is increasing significantly *R. Soc. Open Sci.* **3** 150241
- Ganteaume A, Camia A, Jappiot M, San-Miguel-Ayanz J, Long-Fournel M and Lampin C 2013 A review of the main driving factors of forest fire ignition over Europe *Environ. Manage.* **51** 651–62
- Ganteaume A and Guerra F 2018 Explaining the spatio-seasonal variation of fires by their causes: the case of southeastern France *Appl. Geogr.* **90** 69–81
- Ganteaume A and Jappiot M 2013 What causes large fires in Southern France *For. Ecol. Manage.* **294** 76–85
- Georganos S, Grippa T, Niang Gadiaga A, Linard C, Lennert M, Vanhuyse S, Mboga N, Wolff E and Kalogirou S 2021 Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling *Geocarto Int.* **36** 121–36
- Ghorbanzadeh O, Kamran K V, Blaschke T, Aryal J, Naboureh A, Einali J and Bian J 2019 Spatial prediction of wildfire susceptibility using field survey GPS data and machine learning approaches *Fire* **2** 1–23
- Gonzalez-Olabarria J R, Brotons L, Gritten D, Tudela A and Teres J A 2012 Identifying location and causality of fire ignition hotspots in a Mediterranean region *Int. J. Wildland Fire* **21** 905
- Hagenauer J and Helbich M 2022 A geographically weighted artificial neural network *Int. J. Geogr. Inf. Sci.* **36** 215–35
- Harris C R et al 2020 Array programming with NumPy *Nature* **585** 357–62
- Hengl T, Nussbaum M, Wright M N, Heuvelink G B M and Gräler B 2018 Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables *PeerJ* **6** e5518
- Hodges J L and Lattimer B Y 2019 Wildland fire spread modeling using convolutional neural networks *Fire Technol.* **55** 2115–42
- Jain P, Coogan S C P, Subramanian S G, Crowley M, Taylor S and Flannigan M D 2020 A review of machine learning applications in wildfire science and management *Environ. Rev.* **28** 478–505
- Lagerquist R, Flannigan M D, Wang X and Marshall G A 2017 Automated prediction of extreme fire weather from synoptic patterns in northern Alberta, Canada *Can. J. For. Res.* **47** 1175–83
- Lemaitre G, Nogueira F and Aridas C K 2017 Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning *J. Mach. Learn. Res.* **18** 1–5
- López-Serrano P M, López-Sánchez C A, Álvarez-González J G and García-Gutiérrez J 2016 A comparison of machine learning techniques applied to Landsat-5 TM spectral data for biomass estimation *Can. J. Remote Sens.* **42** 690–705
- Loyola-Gonzalez O 2019 Black-box vs. white-box: understanding their advantages and weaknesses from a practical point of view *IEEE Access* **7** 154096–113
- Lundberg S M and Lee S I 2017 A unified approach to interpreting model predictions *Advances in Neural Information Processing*

- Systems (Long Beach, CA, USA, 4–9 December 2017)* pp 4765–74
- Mayr M J, Vanselow K A and Samimi C 2018 Fire regimes at the arid fringe: a 16-year remote sensing perspective (2000–2016) on the controls of fire activity in Namibia from spatial predictive models *Ecol. Indica* **91** 324–37
- McKinney W 2010 Data structures for statistical computing in python *Proc. 9th Python in Science Conf.* pp 56–61
- Molina J R, Lora A, Prades C, Rodriguez Y and Silva F 2019 Roadside vegetation planning and conservation: new approach to prevent and mitigate wildfires based on fire ignition potential *For. Ecol. Manage.* **444** 163–73
- Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A and Pereira J M C 2012 Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest *For. Ecol. Manage.* **275** 117–29
- Pedregosa F et al 2011 Scikit-learn: machine learning in {P}ython *J. Mach. Learn. Res.* **12** 2825–30
- Penman T D, Bradstock R A and Price O 2013 Modelling the determinants of ignition in the Sydney Basin, Australia: implications for future management *Int. J. Wildland Fire* **22** 469–78
- Pereira A, Pereira J, Libonati R, Oom D, Setzer A, Morelli F, Machado-Silva F and de Carvalho L 2017 Burned area mapping in the Brazilian savanna using a one-class support vector machine trained by active fires *Remote Sens.* **9** 1161
- Riley K L, Grenfell I C, Finney M A and Crookston N L 2014 Utilizing random forests imputation of forest plot data for landscape-level wildfire analyses *Advances in Forest Fire Research ed D X Viegas* (Coimbra: Imprensa da Universidade de Coimbra) pp 583–90
- Riley S J, DeGloria S D and Elliot R 1999 A terrain ruggedness index that quantifies topographic heterogeneity *Int. J. Sci.* **5** 23–27
- Rodrigues M, de la Riva J and Fotheringham S 2014 Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression *Appl. Geogr.* **48** 52–63
- Syphard A D and Keeley J E 2015 Location, timing and extent of wildfire vary by cause of ignition *Int. J. Wildland Fire* **24** 37–47
- Syphard A D, Radeloff V C, Keuler N S, Taylor R S, Hawbaker T J, Stewart S I and Clayton M K 2008 Predicting spatial patterns of fire on a southern California landscape *Int. J. Wildland Fire* **17** 602–13
- Tedim F et al 2022 Forest fire causes and motivations in the southern and South-Eastern Europe through experts' perception and applications to current policies *Forests* **13** 562
- Vecín-Arias D, Castedo-Dorado F, Ordóñez C and Rodríguez-Pérez J R 2016 Biophysical and lightning characteristics drive lightning-induced fire occurrence in the central plateau of the Iberian Peninsula *Agric. For. Meteorol.* **225** 36–47
- Waskom M 2021 seaborn: statistical data visualization *J. Open Source Softw.* **6** 3021
- Zhao Y, Ma J, Li X and Zhang J 2018 Saliency detection and deep learning-based wildfire identification in UAV imagery *Sensors* **18** 712

Topographic and Vegetation Controls on Wildfires in South-East France, 1970-2019

C. Bountzouklis¹, D.M. Fox¹, E. Di Bernardino²

¹University of Côte d'Azur, UMR CNRS 7300 ESPACE, Nice, 06204, France; ²University of Côte d'Azur, UMR CNRS 7351 LJAD, Nice, 06108, France

1. Introduction:

Forest fires in the Euro-Mediterranean zone burn close to 500,000 ha per year. Many of these fires occur in the Wildland Urban Interface (WUI) (Fig. 1) where wildfires pose a serious threat to people and houses. The objective of this study was to quantify temporal changes in environmental factors influencing the spatial distribution of burned area in 3 departments (06, 13, 83) of the SUD-PACA region since 1970.



Fig. 1: Example of Wildland Urban Interface (WUI).

2. Methods:

A burned area polygon database accounting for 3,382 fires and a total burned area of 296,820 ha was obtained from the National Forestry Service (ONF) for fires between 1970 and 2019. Fire contour precision improved with time as satellite imagery became available (Fig. 2).

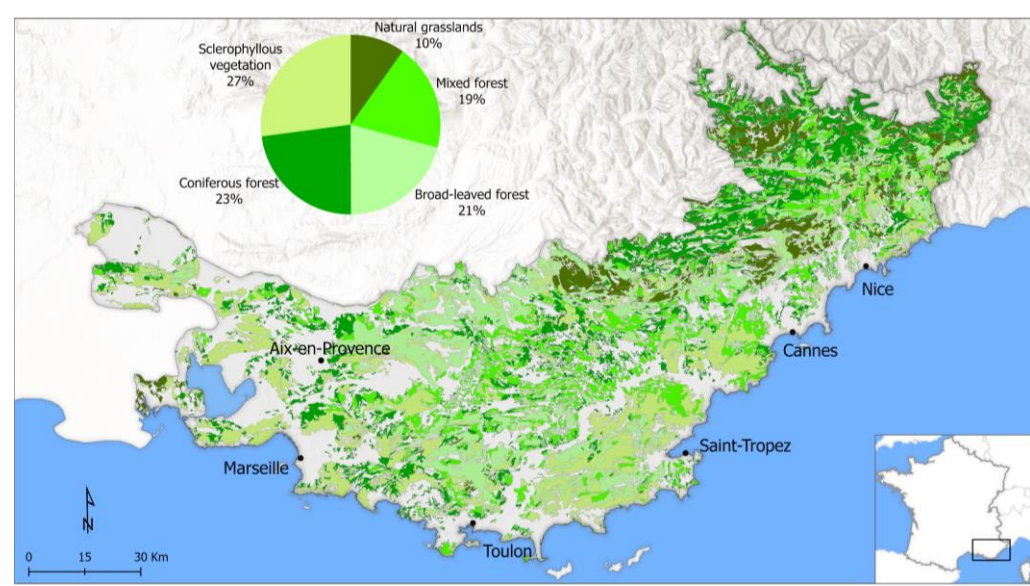


Fig. 3: Corine Land Cover (CLC) database for vegetation mapping.

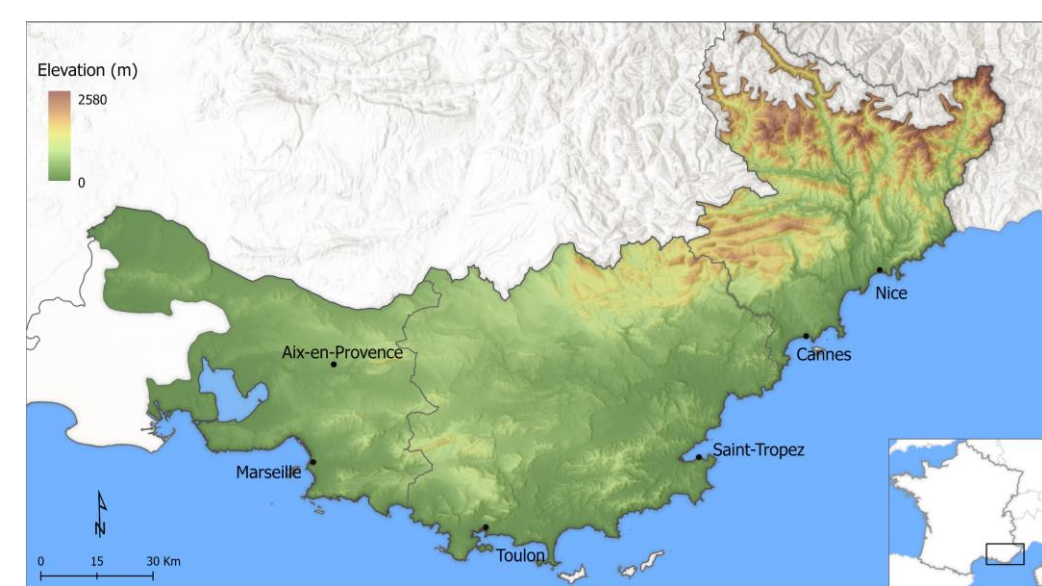


Fig. 4: 5 m Digital Elevation Model (DEM) for slope inclination and aspect.

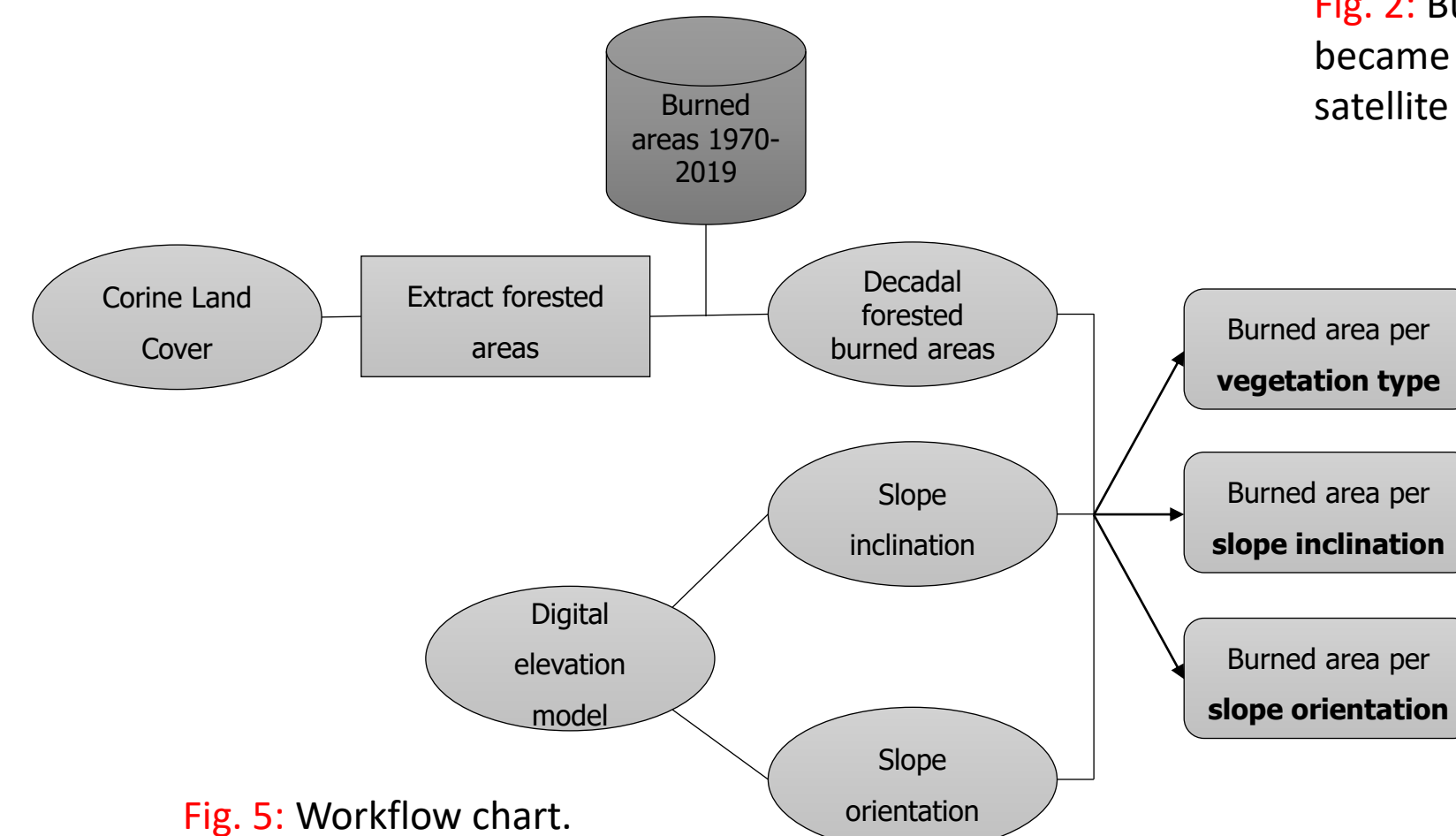


Fig. 5: Workflow chart.

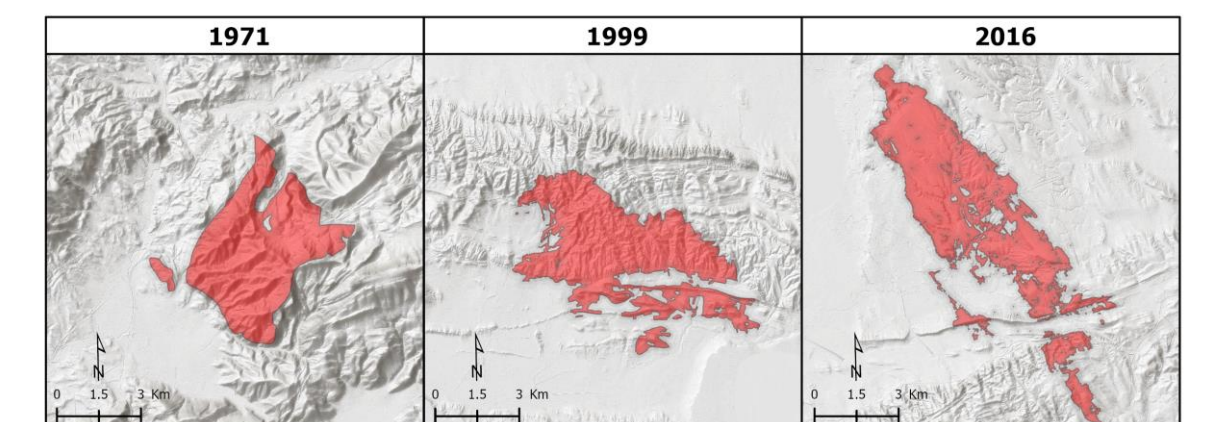


Fig. 2: Burn scar contours became more refined with satellite imagery.

As shown in Fig. 5, burn scar polygons were overlain on slope inclination and aspect to extract percentage burned area per slope and aspect categories. Burned forested area by vegetation type was also extracted in this way.

Data were calculated according to 25-year (1970-1994, 1995-2019) and decadal time intervals.

3. Results and Discussion:

Burned area is highly variable from one year to another, but big fire years tend to disappear after an improvement in fire prevention and suppression starting in the early 1990's; only 2003, the hottest/driest year on record stands out in the 1994-2019 time interval (Fig. 6).

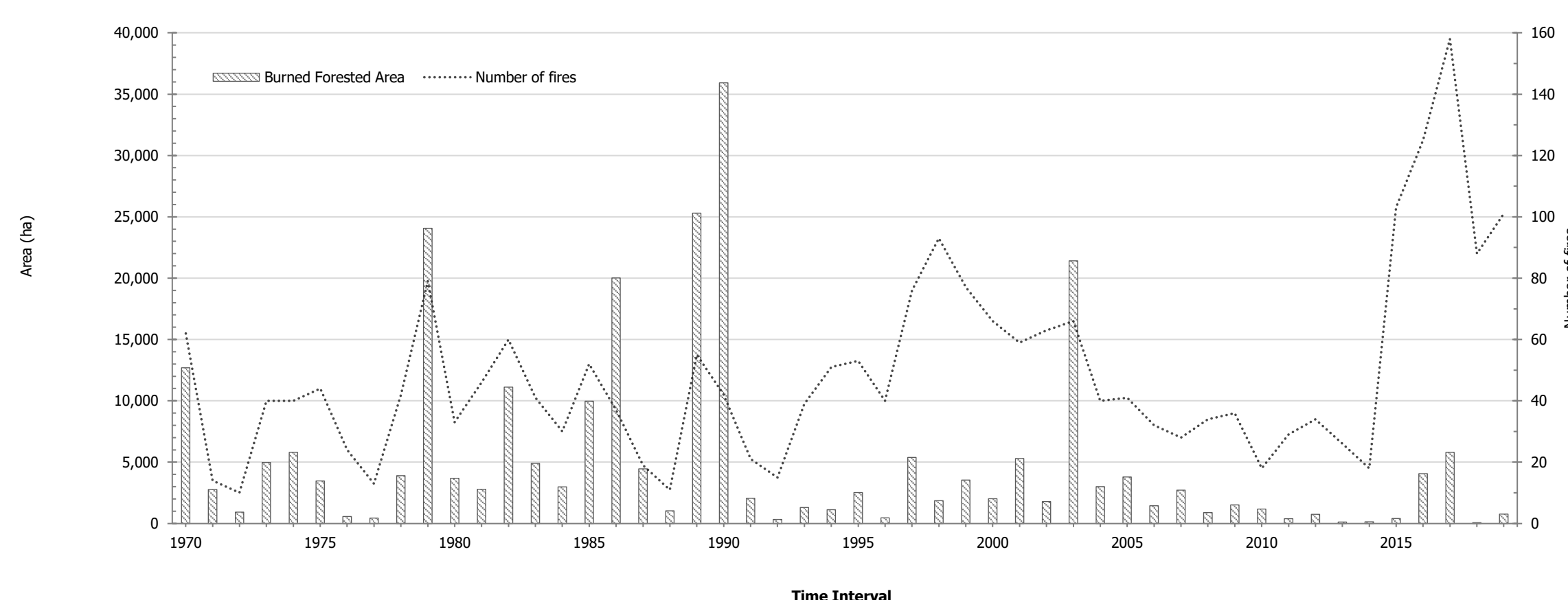


Fig. 6: Annual burned area and number of fires, 1970-2019.

Burned patch sizes shrunk in the second time interval (1995-2019) compared to the first (1970-1994) (Fig. 7). Some persistent hotspots persevere through time in both the Bouches-du-Rhône and Var departments, though both their size and fire frequency diminished.

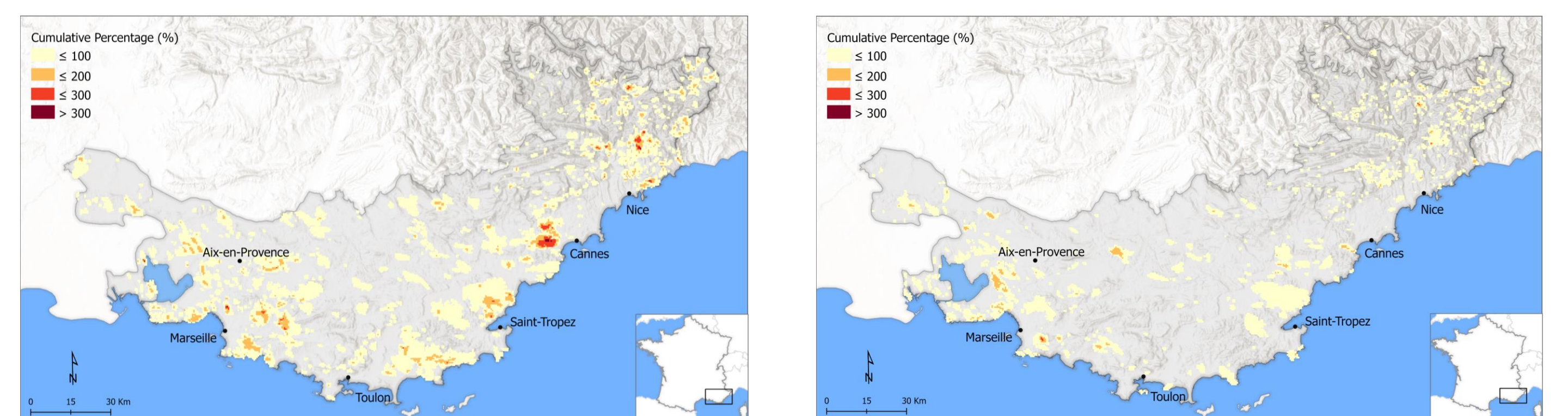


Fig. 7: Burned area frequency for the 1970-1994 (left) and 1995-2019 time intervals (right).

Topographic and vegetation results are plotted as the distribution of total burned area according to category (sum of all bars in a decade=100%) in top row of figures and as a percentage of the category found in the field (i.e.. Percentage of north-facing slopes burned, for example) in bottom row figures. These values are sensitive to total burned area, so bar height decreases after the big fire decades of 1970-1979 & 1980-1989.

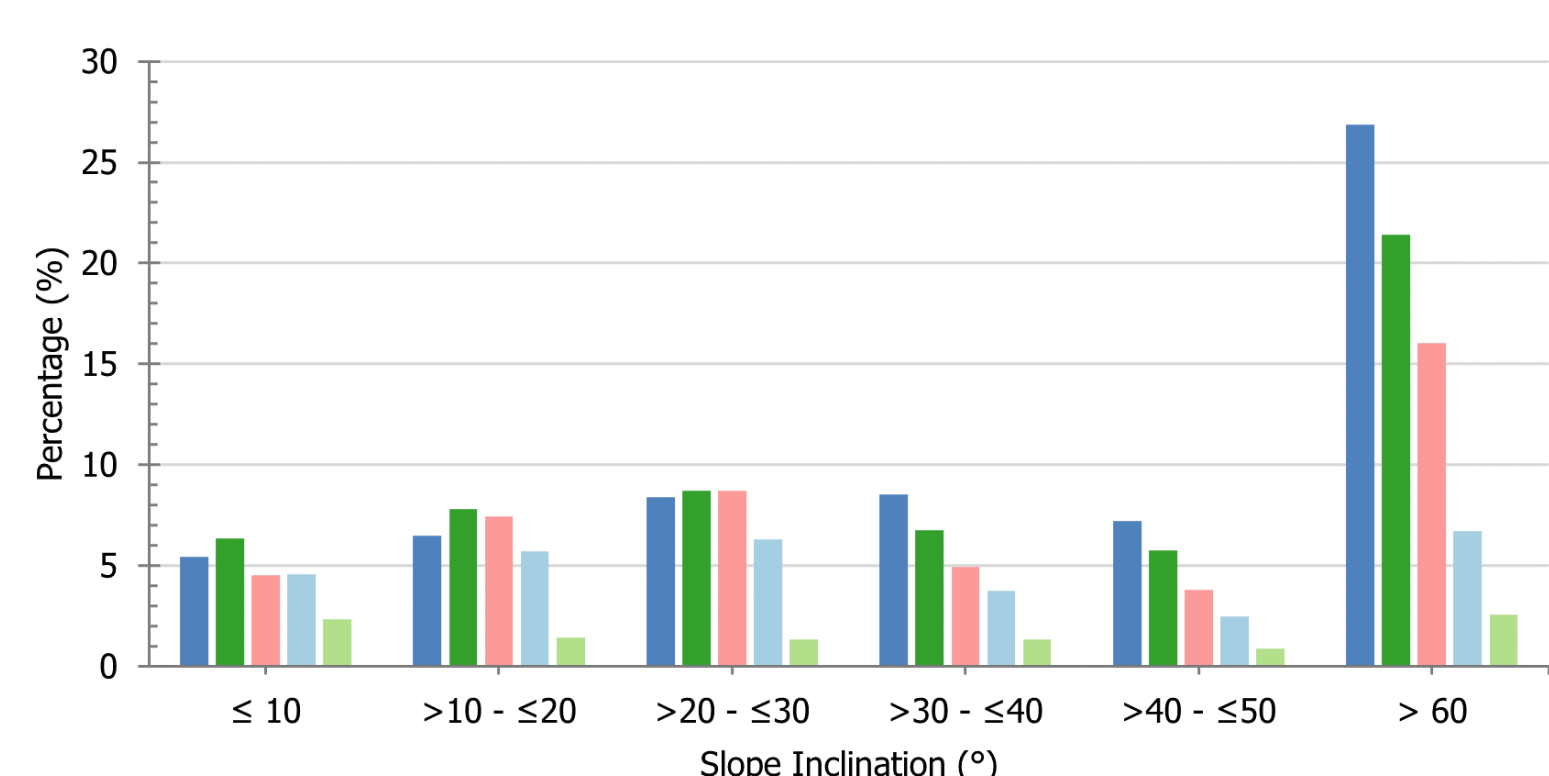
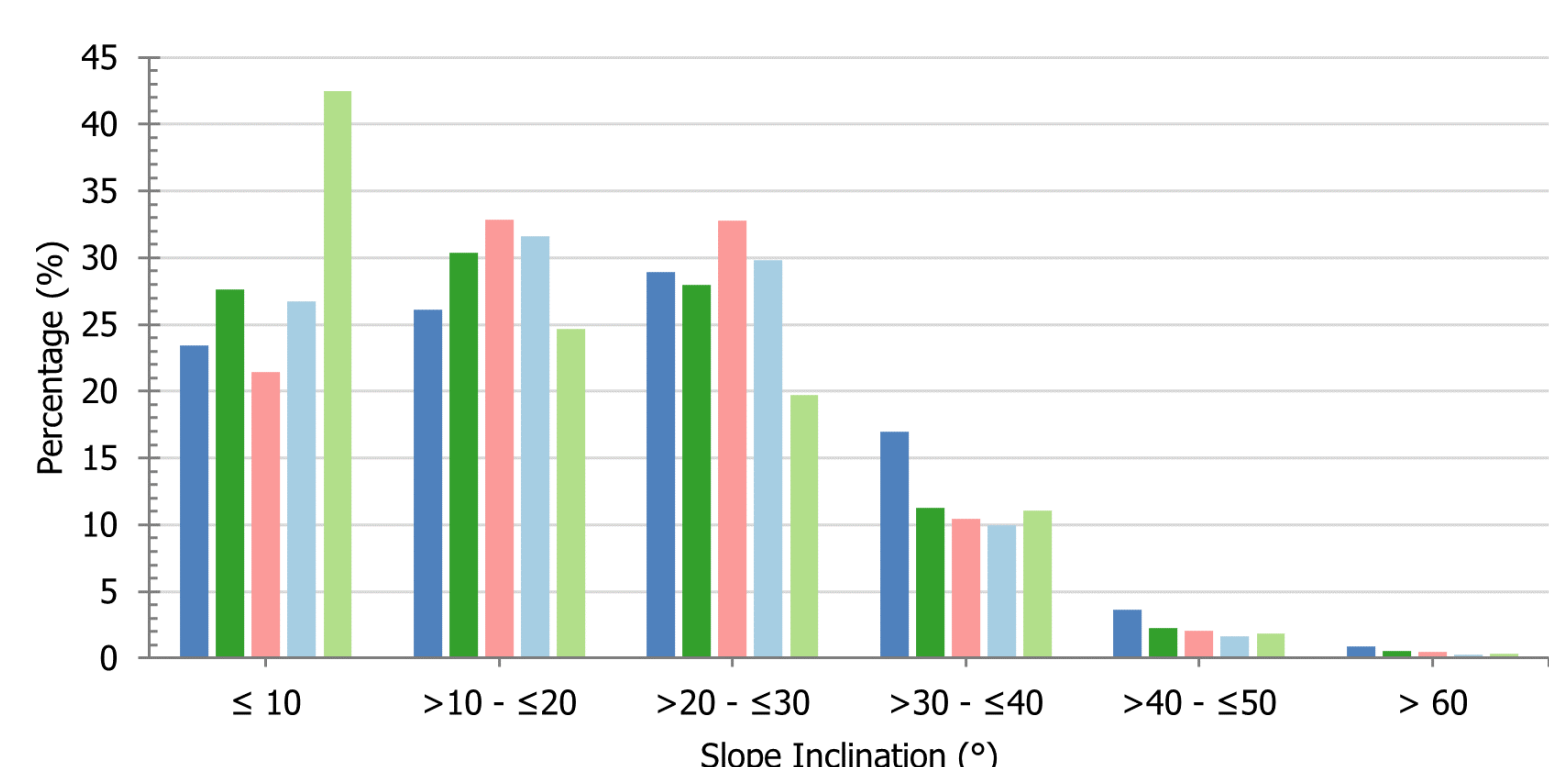


Fig. 8: Burned area per decade expressed as a percentage of total burned area per slope inclination category (sum of bars per decade=100%) (top) and as a percentage of slope category burned (bottom).

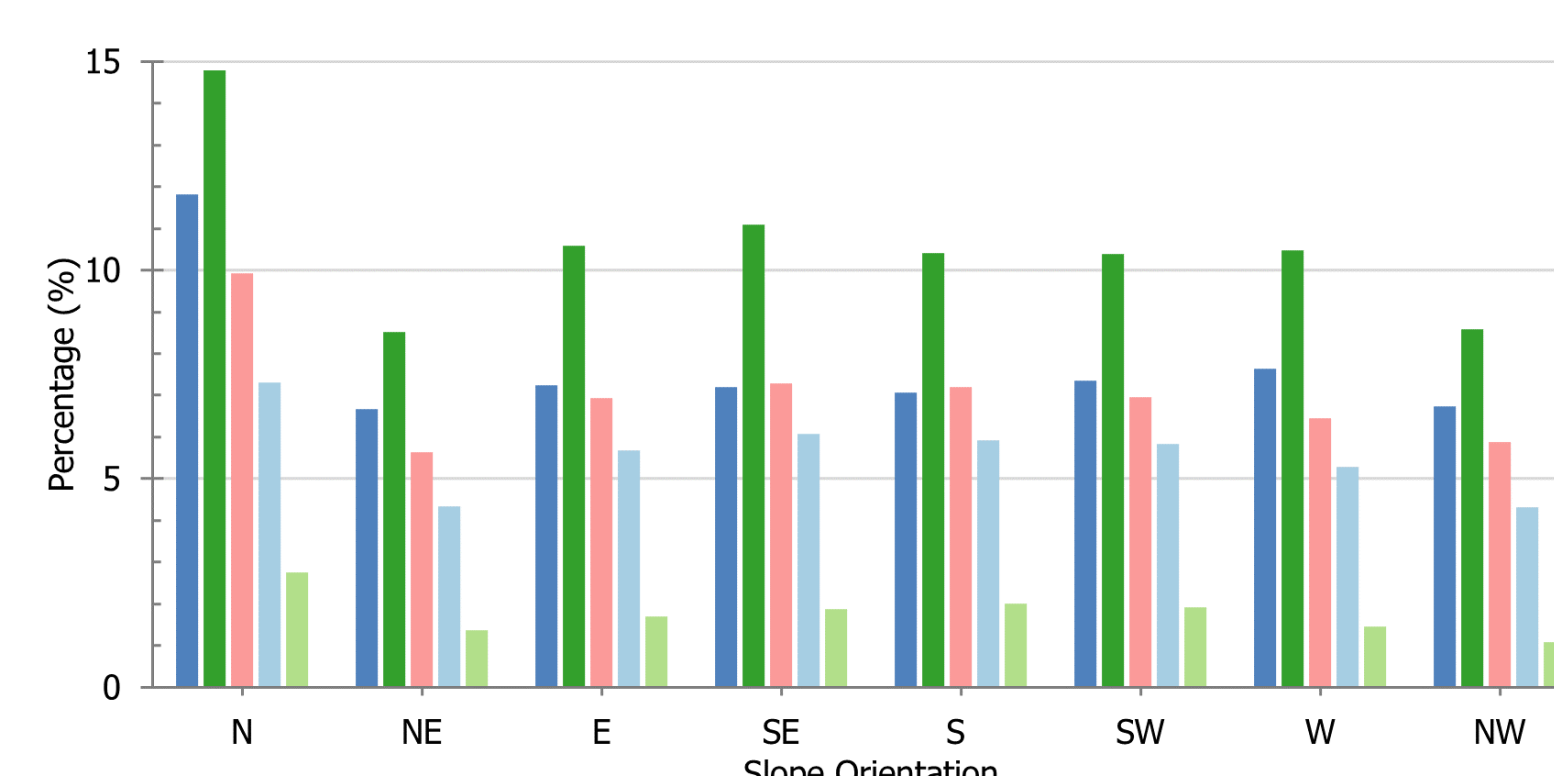
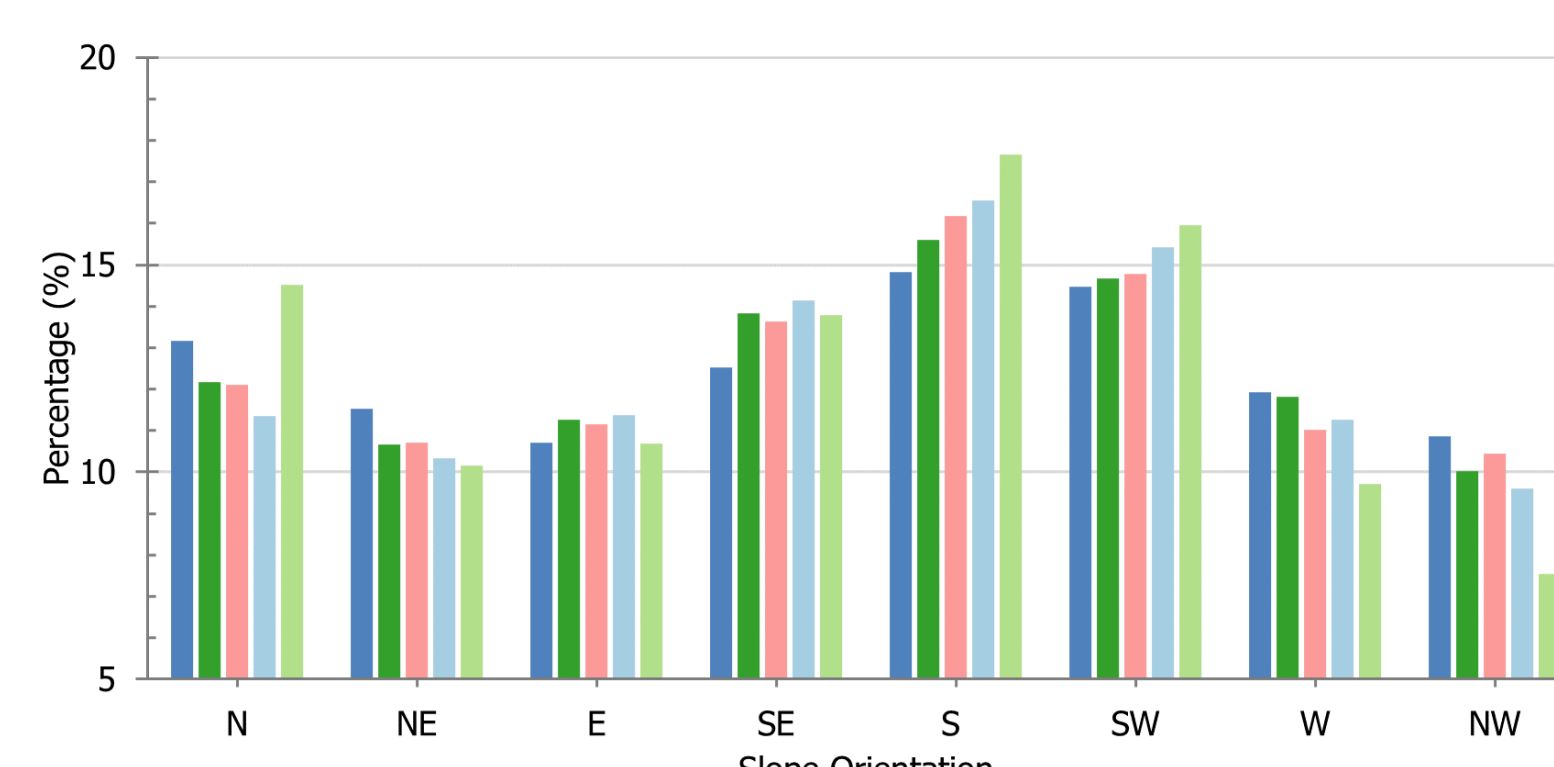


Fig. 9: Burned area per decade expressed as a percentage of total burned area per slope aspect category (sum of bars per decade=100%) (top) and as a percentage of slope aspect burned (bottom).

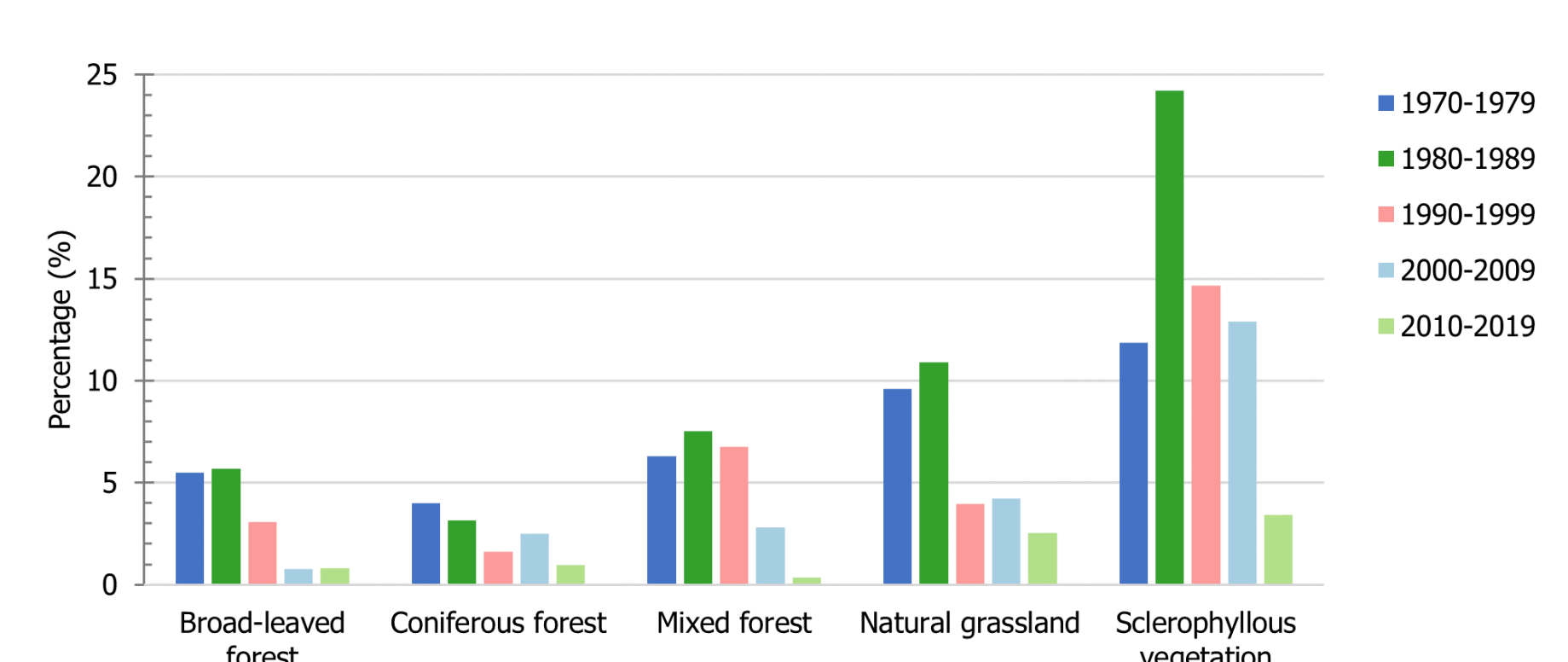
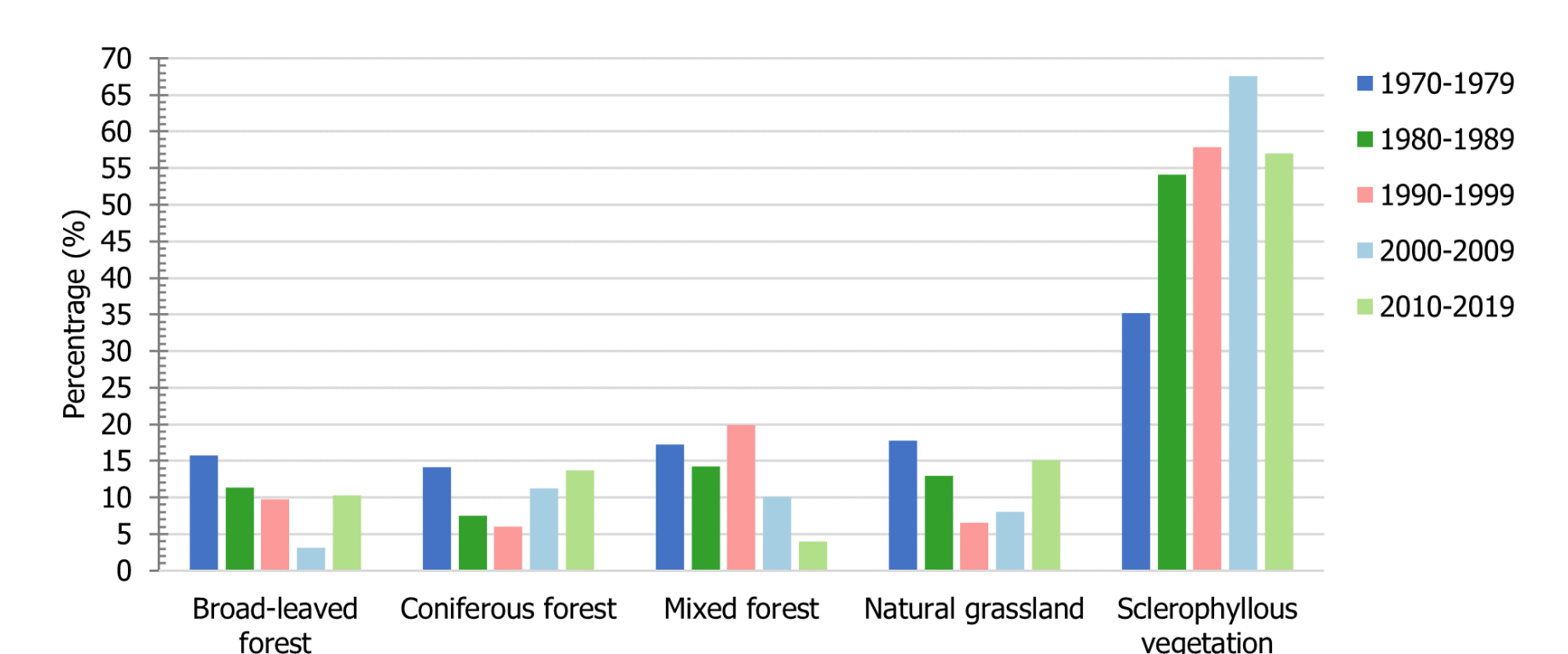


Fig. 10: Burned area per decade expressed as a percentage of total burned area per vegetation category (sum of bars per decade=100%) (top) and as a percentage of vegetation category burned (bottom).

Slope Inclination

Absolute burned area is generally greater on intermediate slopes (Fig. 8 top), and there's a tendency for burned areas to shift towards gentler slopes inclinations after 1990-1999; this coincides with the shift to more effective fire repression and suppression strategies.

The relative burned area (percentage of slopes burned) is greatest for steeper slopes (>60%) as shown in Fig. 8, bottom; this, however, concerns very little burned area. Slope inclinations most susceptible to burning remain the intermediate categories of about 10-40%.

Slope Aspect

Absolute burned area is greater on South-facing slopes (Fig. 9 top). South-facing slopes are drier and have more houses/roads with greater fire ignition risk. In addition, the Sud-PACA region has greater South-facing slope area than North-facing slopes. Surprisingly, a greater percentage of North-facing slopes is burned than for South-facing slopes (Fig. 9 bottom). This unexpected can be linked to the direction of maximum wind speed, which is frequently from the North (Mistral winds).

Vegetation Type

Sclerophyllous vegetation (i.e.. Shrubland) has the greatest absolute burned area (Fig. 10 top). Shrublands are drier and allow greater wind speeds so fires propagate more quickly. This finding is confirmed when burned area is expressed as a percentage of vegetation category area (Fig. 10 bottom).

4. Conclusion

Burned area decreased sharply in the early 1990s when the SDIS firefighting service developed more efficient fire repression and suppression strategies after the catastrophic fires of the 1980s. Frequently burned patches diminished in both areal extent and fire frequency afterwards. Intermediate slopes are the most susceptible to burn, but the study reveals a shift to lower slope inclinations, and this may be due to the rapid suppression in the initial stages of the fire which thereby limits propagation to steeper slopes. Although fires are frequently found on South-facing slopes due to drier conditions and a greater potential for fire ignition, North-facing slopes have greater wind speeds and a greater proportion of these slopes are burned compared to South-facing slopes. Finally, shrublands are the most susceptible to wildfires. The next step is to use machine learning algorithms in conjunction with weather and land cover data to better predict forest fire risk in the future.

5. Acknowledgements

The authors would like to express their gratitude to the SUD-PACA region for funding this project in the context of the PACA-Exploratoire call. They are also grateful to the ONF for making the data available in the framework of a bilateral agreement. This work has been supported by the French government, through the 3iA Côte d'Azur Investments and the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.

Fire source classification of unknown-caused fires based on Random Forests

C. Bountzouklis¹, D.M. Fox¹, E. Di Bernardino²

¹Université Côte d'Azur, UMR CNRS 7300 ESPACE, Nice, 06204, France; ²Université Côte d'Azur, UMR CNRS 7351 LJAD, Nice, 06108, France

1. Introduction:

Approximately 70% of all fires between 1973 and 2020 were ignited by an unknown cause in France. Recording fire causes and studying their spatiotemporal patterns is important for establishing useful fire policies since a better understanding can enhance the efficacy of fire prevention measures. This study examines whether machine-learning methods can be used to identify the origin (arson or non-arson) of unknown-caused fires and evaluate the importance of several environmental and anthropogenic factors in determining the ignition source.

2. Study area:

The study area covers the administrative department of "Bouches-du-Rhône", which ranks second in terms of burned area and fire frequency in mainland France (Fig. 1).

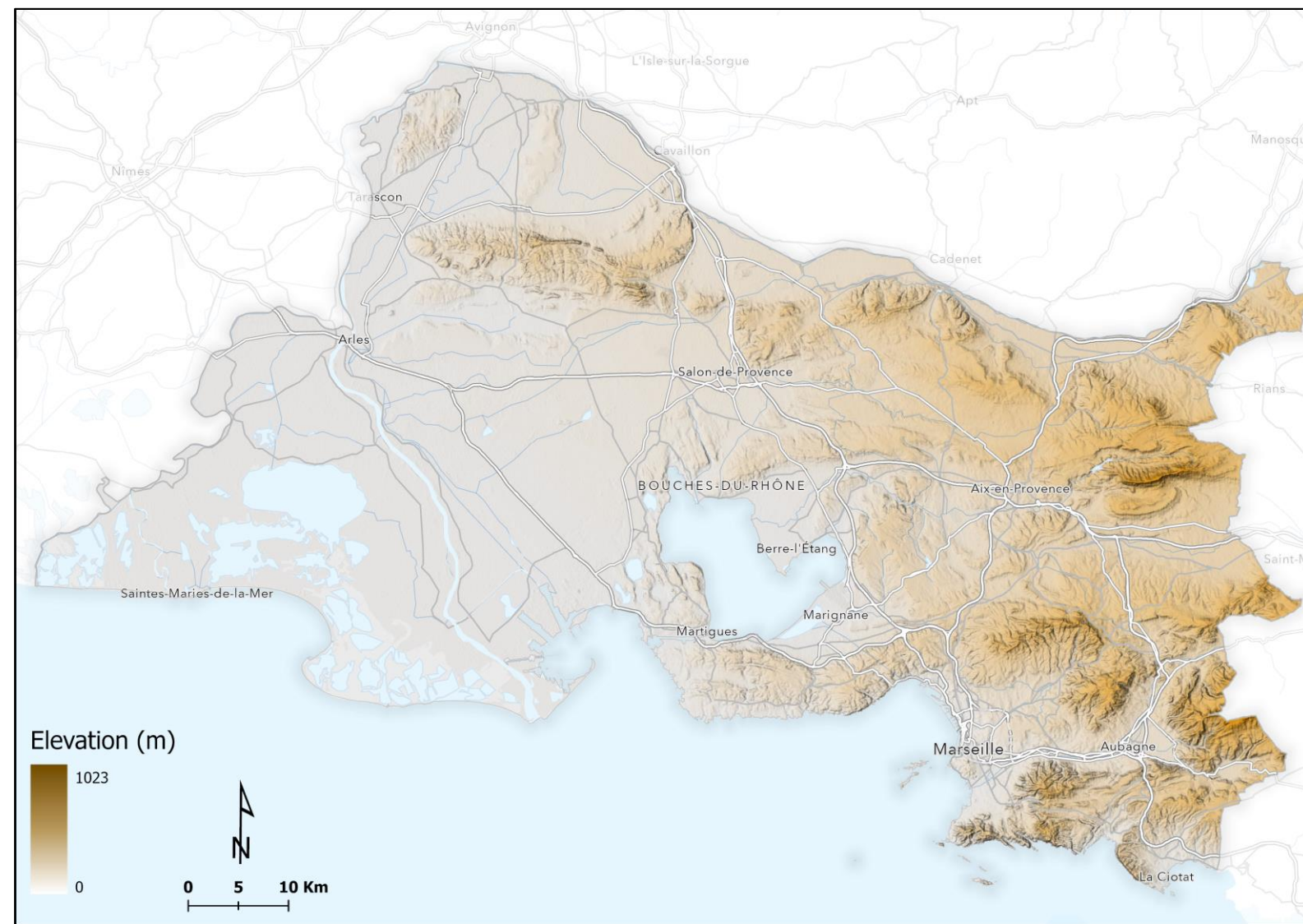


Fig. 1: Fire ignition points within the Departmental limits of Bouches-du-Rhône.

Fires caused by arson account for (Fig. 2):

- 63 % of the total burned area
- 51 % of all fire ignitions
- 60 % of large fires (>100 ha)

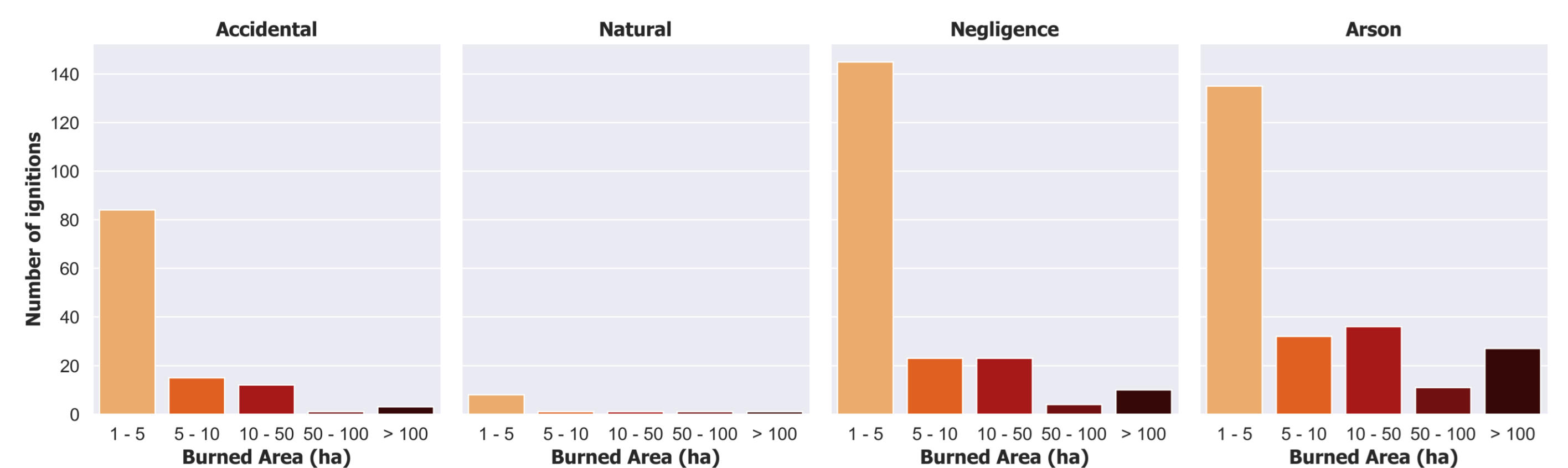


Fig. 2: Number of fires per volume category of burned area (fires > 1ha) and ignition cause from 1973 to 2020.

3. Data & Methods:

A geographic database that contains exact coordinates of fire ignitions was utilized that is provided by the National Forestry Office. Due to the limited recorded number of fires (n=323) caused by accident, negligence and lightning strikes (natural) but also due to the significance of arson fires in the area, fire sources were merged into two major categories, arson and non-arson (Fig. 3).

➢ Multiple environmental and anthropogenic factors (Fig. 4) were used to train a model that can identify the ignition source of a fire.

➢ A circular buffer zone (500 m) was created around each fire ignition point to extract the contextual geographic information.

➢ Based on Random Forests, 70% of the dataset was utilized to train the model, and 30% to test the accuracy in predicting the cause of a fire.

➢ The process of splitting the data (keeping the same ratio) and executing the model was iterated (n=300).

➢ Feature importance scores were calculated using the Gini impurity method.

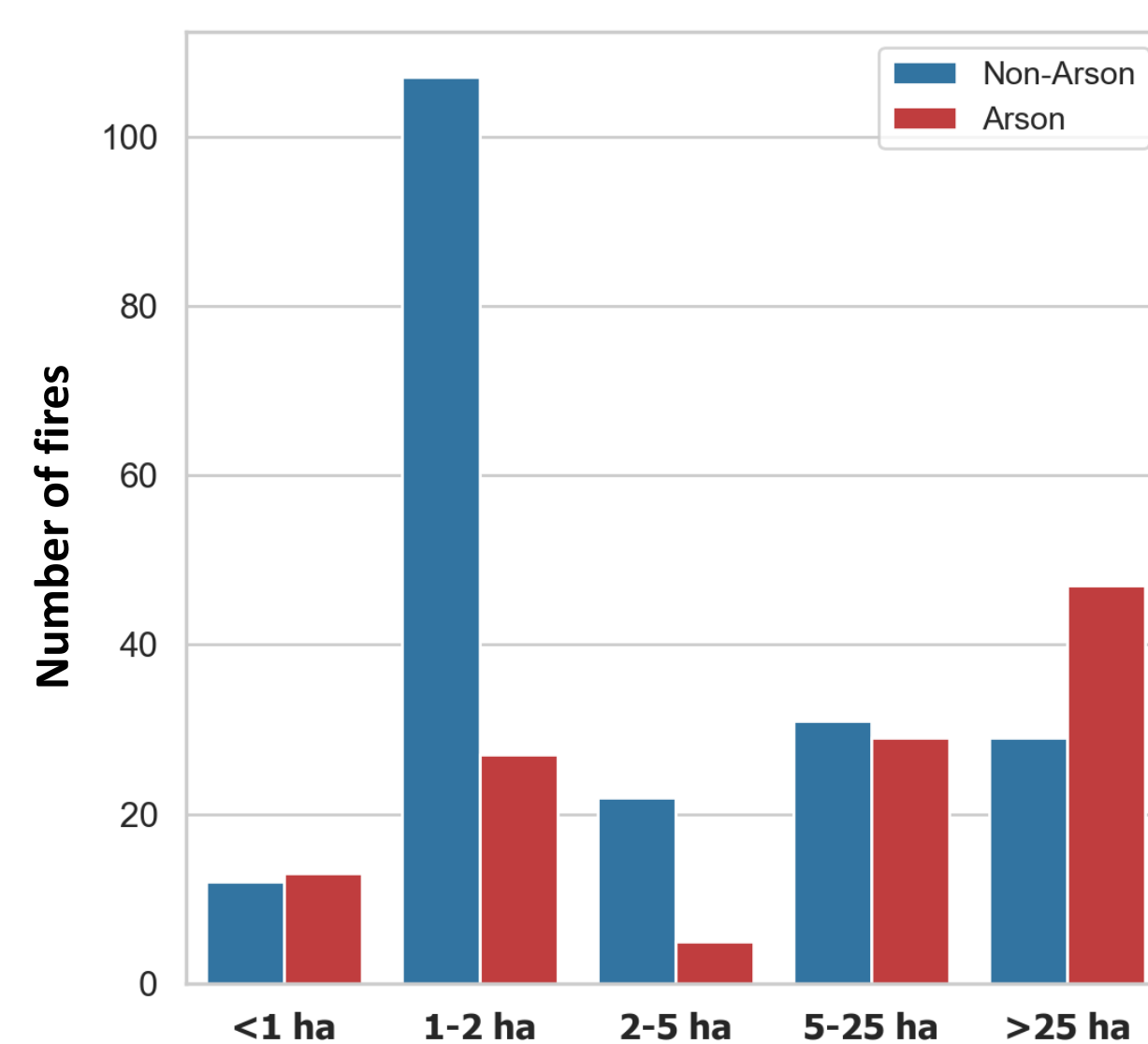


Fig. 3: Number of fires per cause and burned area size.

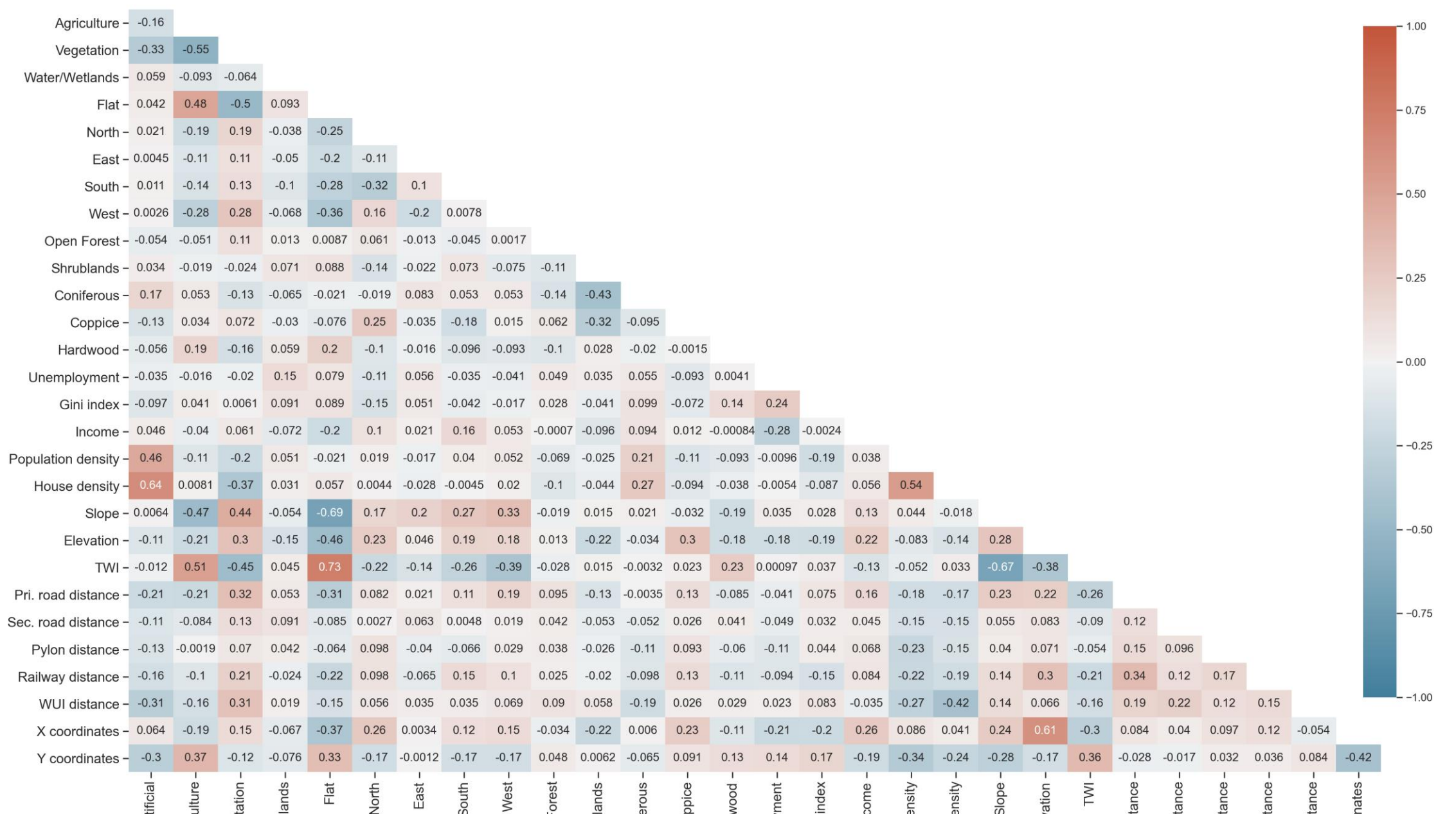


Fig. 4: Kendall's Tau rank correlation coefficient heatmap of explanatory variables.

4. Results:

The accuracy of the model to classify unknown caused fires varies substantially ranging from 56% to 76% (median 67%), due to the small size of the dataset (n=323) (Fig. 5).

➢ Overall, anthropogenic features surpass in importance the rest (Fig. 6).

➢ Topographic factors are more important than land cover and vegetation type while spatiotemporal variables hold the lowest importance.

➢ Secondary road distance, Shrublands and Unemployment rate displaying the highest importance.

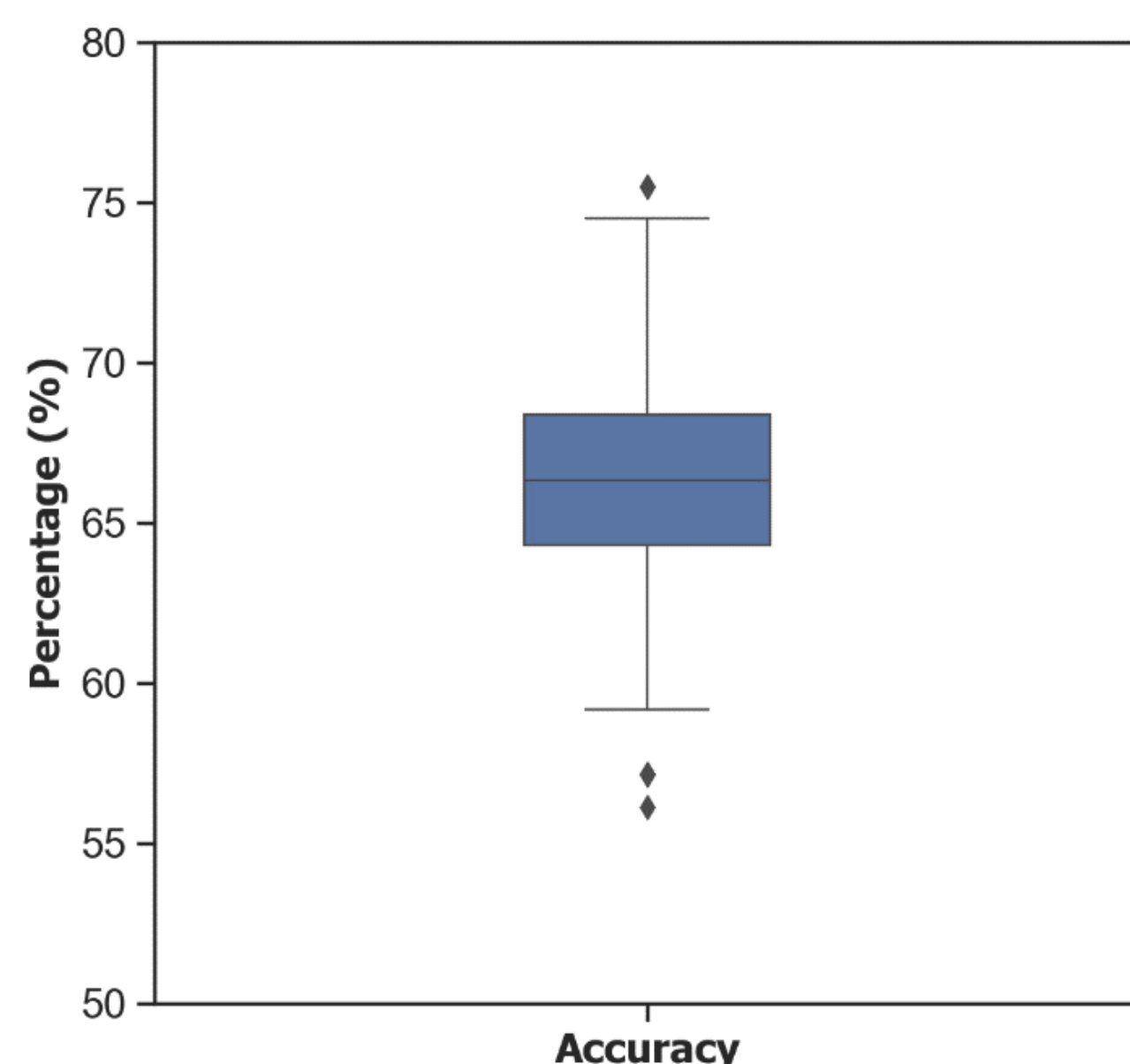


Fig. 5: Boxplot representing Random Forest's classification accuracy for all iterations (n=300).

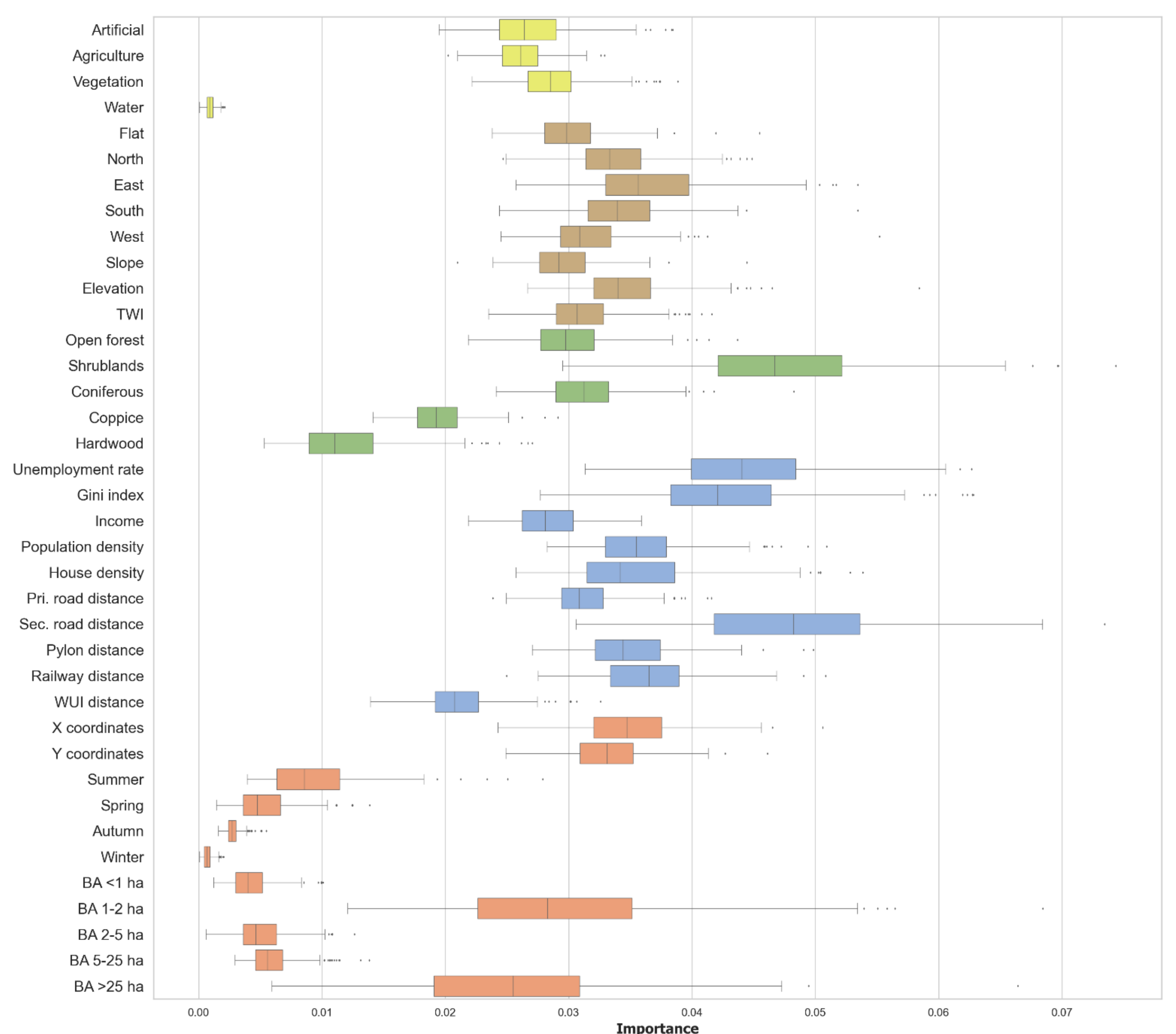


Fig. 6: Distribution of the variable importance values for all iterations (n=300).

4. Conclusion:

The source of unknown caused fires can be identified at an acceptable level of accuracy even with a limited number of fires. Overall performance of such models would greatly benefit from the exploitation of larger datasets as well as from the inclusion of weather-related variables. As location holds particular importance over certain fire causes, spatial extensions of ML algorithms such as Geographic random forests could potentially provide significant enhancements over the original algorithm.

5. Acknowledgements:

The authors would like to express their gratitude to the SUD-PACA region for funding this project in the context of the PACA-Exploratoire call. They are also grateful to the ONF for making the data available in the framework of a bilateral agreement. This work has been supported by the French government, through the 3iA Côte d'Azur Investments and the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.