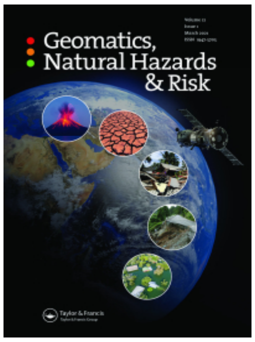

Geomatics, Natural Hazards and Risk

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgnh20

A spatially explicit containment modelling approach for escaped wildfires in a Mediterranean climate using machine learning

Gbenga Lawrence Alawode, Pere Joan Gelabert & Marcos Rodrigues

To cite this article: Gbenga Lawrence Alawode, Pere Joan Gelabert & Marcos Rodrigues (2025) A spatially explicit containment modelling approach for escaped wildfires in a Mediterranean climate using machine learning, Geomatics, Natural Hazards and Risk, 16:1, 2447514, DOI: 10.1080/19475705.2024.2447514

To link to this article: https://doi.org/10.1080/19475705.2024.2447514

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 11 Jan 2025.

Submit your article to this journal

Article views: 249

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tgnh20

GEOMATICS, NATURAL HAZARDS AND RISK
2025, VOL. 16, NO. 1, 2447514
https://doi.org/10.1080/19475705.2024.2447514

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

A spatially explicit containment modelling approach for escaped wildfires in a Mediterranean climate using machine learning

Gbenga Lawrence Alawode[a,b*], Pere Joan Gelabert[a†] and Marcos Rodrigues[c‡]

[a]Department of Agricultural and Forest Sciences and Engineering (DCEFA), Universitat de Lleida, Lleida, Spain; [b]Faculty of Science, Forestry and Technology, University of Eastern Finland, Joensuu, Finland; [c]Department of Geography, University of Zaragoza, Zaragoza, Spain

ABSTRACT

Wildfires are particularly prevalent in the Mediterranean, being expected to increase in frequency due to the expected increase in regional temperatures and decrease in precipitation. Effectively suppressing large wildfires requires a thorough understanding of containment opportunities across landscapes, to which empirical spatial modelling can contribute largely. The previous containment model in Catalonia failed to account for the crucial roles of weather conditions, lacked temporal prediction and could not forecast windows for containment opportunities, prompting this research. We employed a detailed geospatial approach to assess the spatial-temporal variations in containment probability for escaped wildfires in Catalonia. Using machine learning algorithms, geospatial data, and 124 historical wildfire perimeters from 2000 to 2015, we developed a predictive model with high accuracy (Area Under the Receiver Operating Characteristics Curve $= 0.81 \pm 0.03$) over 32,108 km^2 at a 30-meter resolution. Our analysis identified agricultural plains near non-burnable barriers, such as major road corridors, as having the highest containment probability. Conversely, steep mountainous regions with limited accessibility exhibited lower containment success rates. We also found temperature and windspeed to be critical factors influencing containment success. These findings inform optimal firefighting resource allocation and contribute to strategic fuel management initiatives to enhance firefighting operations.

ARTICLE HISTORY
Received 2 July 2024
Accepted 22 December 2024

KEYWORDS
Megafires; random forest; geospatial; fire suppression; spatial-temporal

CONTACT Marcos Rodrigues rmarcos@unizar.es
*University of Nevada, Graduate Program of Hydrologic Sciences, Reno, Nevada, USA
†Joint Research Unit CTFC-Agrotecnio-Cerca, Solsona, Spain.
‡GEOFOREST Research Group, University Institute for Environmental Sciences (IUCA), Zaragoza, Spain.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

1. Introduction

Wildfires are a prevalent disturbance in the Mediterranean region, playing a crucial role in shaping its ecosystems (O'Hara 2014; Gonçalves and Sousa 2017). They have also been historically utilized as a tool for forest management (Badia et al. 2019). Spain, among the European countries in the Mediterranean, experiences some of the highest incidences of wildfires and burned areas (Martínez et al. 2009; European Commission: Directorate-General for Environment et al. 2016), resulting in significant environmental degradation, economic losses, and human costs (Viana-Soto et al. 2017). In fact, Spain ranks second in fire frequency for wildfires larger than one hectare and highest in total burned area (Rodrigues et al. 2020a). The impacts of wildfires extend beyond direct consequences, such as changes in flora and fauna within affected ecosystems. They also give rise to indirect repercussions, disturbing both the economic flows and environmental services of the fire-affected areas and their surrounding regions (Gonçalves and Sousa 2017). Furthermore, the tangible losses incurred during wildfire events can be measured in monetary terms, reflecting their significant financial impact. Additionally, there are intangible losses, such as ecological damage and water catchment degradation, that are challenging to quantify but nonetheless exert profound effects on the environment and the economy (Handmer and Proudley 2008).

In recent years, there has been a notable increase in the risk of wildfires in peri-urban areas and fragmented housing situated in semi-natural regions across the Mediterranean (Alcasena et al. 2018). In Catalonia, this trend is exacerbated by a rise in the frequency of extreme weather events, coupled with fuel buildup resulting from land abandonment (Gelabert et al. 2022). These factors have contributed to a rise in the occurrence of large-scale wildfires (González and Pukkala 2007). However, the implementation of comprehensive prevention plans to mitigate the adverse impacts of these wildfires is hindered by limited funding and the predominantly private and fragmented ownership of forested land. With over 77% of forested areas owned by numerous individuals, often with small property sizes averaging around 30 hectares, coordinating large-scale prevention efforts poses a significant challenge (Gonzalez-Olabarria et al. 2019). Despite these challenges, there is a concerted effort to minimize the negative effects of wildfires in this fire-prone landscape (Gill et al. 2013).

Suppressing extreme wildfires is considerably more challenging compared to the initial attack of smaller wildfires. Small wildfires often occur within a more confined fire environment, characterized by limited fuel, mild weather conditions, and less rugged terrain. Consequently, firefighting resources allocated to these incidents are comparatively limited. It is widely recognized that under extreme weather conditions, wildfires' behaviour often surpasses suppression efforts' capabilities. Once a wildfire gains momentum, the effectiveness of suppression measures diminishes significantly, particularly when ample fuel is available (San-Miguel-Ayanz et al. 2013; Fernandes et al. 2016).

The influence of weather conditions, including temperature and windspeed, on wildfire behaviour has been widely documented in the literature (Finney et al. 2009; Díaz-Avalos et al. 2016; Gonzalez-Olabarria et al. 2019; Moreno et al. 2023). In Mediterranean regions, most burned areas are attributed to large wildfire events resulting from the complex interplay of various weather processes across different spatial and temporal scales (Pereira et al. 2005). As regional climate projections indicate warmer and drier climates across the region, the Mediterranean is expected to experience summer warming ranging from 1.22 to 8.49 °C and a significant decline in precipitation of 16% to 49% by the end of the century (Cos et al. 2022), resulting in an increased likelihood of large wildfires, particularly during wildfire seasons (Calheiros et al. 2021; Galizia et al. 2023). This is apparent in the lengthening of wildfire seasons and the occurrence of large wildfires outside the summertime (Figure S8). As a result, models aimed at describing wildfire behaviour and predicting suppression capability must adequately account for the crucial influence of weather variables in conjunction with other essential factors impacting these phenomena.

There are different methods used for developing models that describe wildfire behaviour; however, recent analytical trend has shifted towards machine learning algorithms due to their capability to produce models that address different classification and regression problems. Their use cut across several fields of study, including environmental and fire science. Amongst the various machine learning algorithms, many researchers in the Mediterranean climate, especially in Catalonia (Oliveira et al. 2012; Rodrigues et al. 2019, 2020b; Gelabert et al. 2022), have demonstrated the immense capability of the Random Forest algorithm, recommending its adoption for this research study. The existing containment model in Catalonia (Rodrigues et al. 2020b) faced a significant limitation: it disregards the impact of weather conditions and ignores the temporal dimension, hence unable to forecast windows of opportunity for containment. The aforementioned limitation arises from constraints related to model robustness and the systematic approach to accurately represent changes in weather conditions during a wildfire event that occur for several days. Furthermore, in a study conducted in the northern Rocky Mountains, a similar limitation was identified, as the model failed to account for the influence of variations in fire-weather conditions (O'Connor et al. 2017). Additionally, a containment model developed by Finney et al. (2009) employed a generalized linear mixed model that incorporated factors such as fire size and fuel type, yet it did not adequately address the impact of weather variables. As a result, we have designed our study to address these shortcomings effectively. We hypothesize that considering the impact of weather variables can enhance the practical utility of our model. This is because the weather conditions can influence the behaviour of other incorporated variables. Notably, windy conditions may boost wildfire spread, while severe winds can hamper the effectiveness of airborne support. Additionally, intense temperatures may increase evaporation rates in river channels, thereby reducing water availability and potentially compromising their effectiveness as firebreaks. Moreover, fuel moisture content, which directly affects wildfire intensity, spread, and suppression efforts, is closely tied to temperature.

Although the currently available containment model demonstrates significant predictive accuracy, integrating weather parameters will strengthen its practicality and applicability. Furthermore, we can introduce a temporal dimension alongside the spatial dimension to account for weather variations throughout wildfire seasons. Therefore, our study aims to develop a robust yet concise machine learning model for predicting wildfire containment in Catalonia. This model will consider the interaction of weather conditions and various biophysical and environmental factors at a fine-scale resolution. We also seek to capture the temporal dynamics of containment, focusing on Catalonia's wildfire seasons. By doing so, we aim to generate probability maps that illustrate containment success across Catalonia's diverse landscapes. Additionally, we will evaluate the predictive performance of different classification and regression algorithms, identifying key variables that significantly influence containment outcomes.

2. Materials and method

2.1. Overview of the workflow

This study followed a modelling approach to calibrate presence-absence models using machine learning techniques, specifically Random Forest (Figure 1). To this effect, we defined 'presence' as areas where wildfire containment efforts have succeeded and 'absence' as the inner burned areas for each wildfire perimeter. The final Random Forest model was constructed in a series of steps, broadly sub-divided into four stages: Firstly, we obtained all the wildfire perimeter files of Catalonia for the study period and selected the wildfires with burned areas meeting our threshold for escaped wildfires. Secondly, we developed a set of predictors relating to weather conditions,

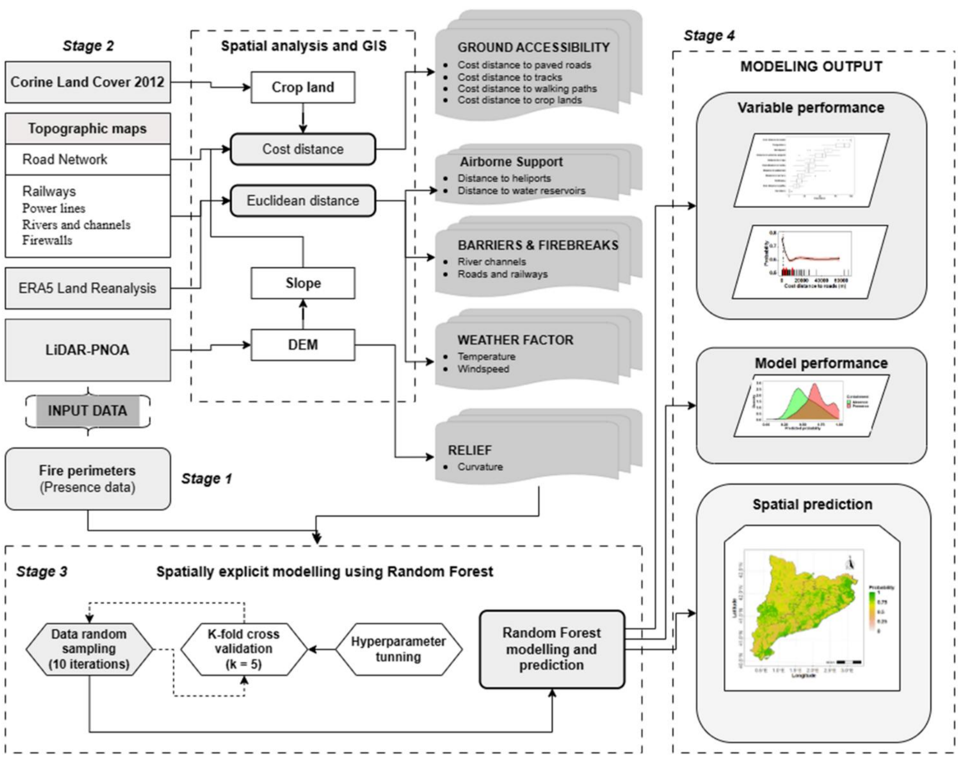

Figure 1. Methodological workflow for modelling fire containment probability in Catalonia. We provided the containment probability at a 30-m spatial resolution for the fire seasons (May–September).

airborne support, wildfire spread obstacles, and surface accessibility. We then built several RF models by resampling and bootstrapping the absence and presence data to ensure a random error distribution and avoid bias in selecting observations. We also fitted several logistic regression and support vector machine models and tested the performance against our RF models. The final step involved model evaluation using the test sample and generating a fine-resolution raster of containment probability across the landscape.

2.2. Study area

Catalonia is one of the seventeen autonomous regions in Spain. Located in the Northeastern part of Spain, it borders the Mediterranean Sea in the eastern part, the autonomous regions of Aragon in the west and Comunidad Valenciana to the south, Andorra and France to the north along the Pyrenees mountains. The area spans 32,108 km^2 and is divided into four provinces, including Barcelona (serving as the capital), Lleida, Tarragona, and Girona (Figure 2(C)). Demographically, Catalonia has an estimated population of 8 million inhabitants (IDESCAT 2023), of which 41% are found in the metropolitan area of Barcelona (AMB 2023). Its characteristic Mediterranean climate results in mild winters and warm and dry summers. This makes the region highly prone to Wildfire (Díaz-Delgado et al. 2004a), but a high diversity in temperature and rainfall patterns across the territories creates a high diversity in species, landscape characteristics and wildfire regimes (Díaz-Delgado et al. 2004b; Castellnou et al. 2009). There is a noticeable variation in mean annual rainfall across different provinces in Catalonia. The southern areas, including parts of Barcelona, Lleida, and significant portions of Tarragona, experience relatively dry conditions, with an annual rainfall of < 700 mm. In contrast, the

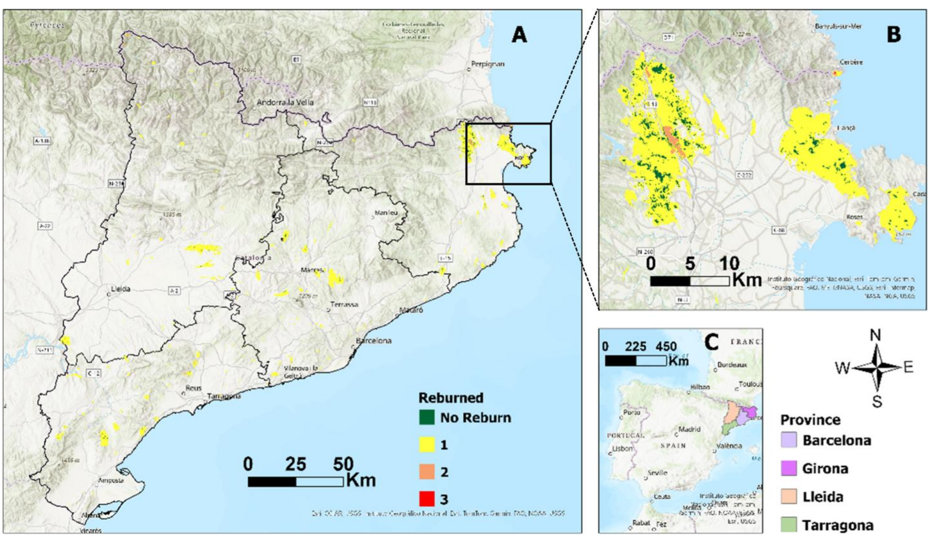

Figure 2. Study location of Catalonia with the fire reburned times between 2000 to 2015 (A); a close-up view of the fire scar in Girona (B); the provinces in Catalonia and their location within Spain (C).

Pyrenees Mountain range in the northern region receives the highest precipitation levels, with mean annual rainfall exceeding 1200 mm. The provinces of Girona and the northern areas of Barcelona and Lleida constitute the region's wetter zones, with annual rainfall equal to or greater than 700 mm (Lopez-Bustins et al. 2020). Temperatures vary throughout the year, with a mean yearly temperature of 16.17 °C, 0.44% lower than Spain's average. Additionally, the wind pattern across the region varies considerably, with the province of Girona having significantly greater windy conditions, especially along the coastal areas and Mountainous regions that experience a strong Tramontana wind that can reach an average speed of 40-50 km/hr.

Characteristically, Catalonia is among the most fire-prone regions in the Mediterranean and is composed of many vegetation types, landscape characteristics, climatic variations, wildfire ignition, propagation, and spread patterns. Human-caused ignitions dominate Catalonia and are prevalent around roads and urban areas (Gonzalez-Olabarria et al. 2012, 2015), although there are a few lightning-induced wildfires that are mostly concentrated in the remote areas of southwest and mainland Catalonia (Pineda and Rigo 2017). Considering the wildfire extent, the region's midland hinterland, southwest and northeastern end have a higher burned area concentration (Alcasena et al. 2019) (Figure 2(A,B)), with small wildfires primarily found along the coastline (Rodrigues et al. 2020a). Much of the burned area, accounting for over 70% of the total, resulted from a few notable instances of extreme, large, and sometimes concurrent wildfire events, notably in 1985, 1994, 1998, 2012, and 2019 (source: MAPAMA, n.d.).

2.3. Methods

2.3.1. Response variable

Predicting the success of wildfire containment necessitates a binary response variable of containment success (1) and containment failure (0). We used the obtained wildfire perimeter to generate this binary response variable, where the wildfire perimeter was designated as areas where past wildfires have been successfully contained, while the rest of the inner burned areas were labelled as places with failed containment efforts (O'Connor et al. 2017). We obtained 124 wildfire perimeters corresponding to those wildfires meeting the specified threshold for escaped wildfires from 2000 to 2015 (information on the frequency, total burned area and burned area distribution can be found in Figures S5, S6 and S7, respectively). Despite the availability of different thresholds for escaped wildfires (Arienti et al. 2006; Plucinski 2012), we selected wildfires with burned areas $\geq$ 50 ha following the provision of Rodrigues et al. (2019). This ensured that we obtained enough representative samples while avoiding spatial autocorrelation. We delineated the wildfire perimeter using the base map of Catalonia Forest fires acquired from the *Generalitat de Catalunya* website of the Department of Climate Actions, Food and Rural agenda (GENCAT 2024). From these perimeters, we obtained the presence points ($n = 1424$) with a regular spacing of 500 m and a buffer of 100 m outside of the wildfire perimeter. A similar procedure was used to extract the absence points ($n = 1326$), specifying a 400 m distance between points and a 200 m buffer away from the wildfire perimeter to avoid co-registration with the presence observations (Figure S8).

2.4. Explanatory variables

We initially identified 19 predictor variables based on the literature review (Petrovic and Carlson 2012; Silva et al. 2014; Fernandes et al. 2016; Thompson et al. 2016; O'Connor et al. 2017; Scott et al. 2017; Riley et al. 2018), which corresponded to several aspects of weather conditions, vegetation impedance, airborne support, terrain and ground accessibility, and relief curvature. These variables were then subjected to a series of tests to achieve three principal aims: (i) while the Random Forest algorithm is designed to handle multicollinearity among variables, highly correlated values can still influence the outcome of our prediction; therefore, we ensured that highly correlated values were excluded from the analysis to avoid bias and redundancy, (ii) we focused on developing a parsimonious model, obtaining a high model accuracy with the least number of predictors, (iii) all variables played a significant role in the model, and the removal of that variable will significantly degrade model accuracy. As a first step in this series of tests, we conducted a correlation analysis between variables to remove one of two highly correlated variables (correlation ≥ 0.5) (Figure S1). Furthermore, variables related to vegetation impedance were removed due to circular logic issues since these variables were derived from LiDAR data taken after some of the wildfires occurred and can lead to erroneous results. Discarding those wildfires would hinder weather variability, hence we opted to retain all wildfires and discard the predictor. Finally, variables of lower importance (relative importance $<$ 25%) were removed at the initial stage of model development. After concluding the pre-processing, 11 variables adequately fulfilled our criteria and were included in the model (Figure S3). To provide a degree of comparison and ensure the accuracy of our model, we built other models using the excluded variables and explored the behaviour of covariates and prediction accuracy.

2.4.1. Meteorological factors

Variations in weather conditions, primarily periods of extreme weather conditions, can increase the potential of wildfire escaping the initial attack (Finney et al. 2009; Rodrigues et al. 2019). Besides, wildfire propagation and containment success are opportunistic as wildfire intensity and spread increase during intense weather conditions. Four meteorological variables (Temperature, Windspeed, Relative Humidity, and Precipitation) relating to wildfire spread and intensity were obtained from the ERA5 Land Reanalysis dataset (C3S 2019) and incorporated into the model on a daily temporal scale to account for weather effects (Table S1). These four variables were subsequently reduced to two based on the criteria described in Section 2.4, retaining temperature and wind speed. The major constraint in effectively allocating weather to wildfire location lies in determining the date of each random point corresponding to the weather variable since wildfire spreads at different times, and each wildfire perimeter differs in burn duration and spatial extent. This problem only applies to the inner burned areas, where containment efforts were unsuccessful, as the points in the perimeter all correspond to the documented extinction date of each wildfire event. We, therefore, developed a method to overcome this problem and objectively allocate daily variables to each point by obtaining the ignition point (EGIF 2024) for each wildfire perimeter and building an Euclidean distance from the ignition point over the wildfire perimeter, assuming a constant spread rate. Afterwards, we extracted the distance corresponding to each random point and calculated the date of the wildfire event based on the distance from the ignition point and the size of the wildfire perimeter (referenced calculation in Figure S2). This estimated date was allocated to each absence point. Since the perimeter corresponds to successful wildfire containment, documented extinction dates were assigned to the presence points. After assigning the dates to each presence and absence point, the weather variables corresponding to the new dates were extracted into a database to obtain the weather conditions corresponding to the escaped and contained wildfire points.

2.4.2. Ground accessibility

We obtained four vector layers corresponding to the entire road networks from the National Topographic Database (BTN25) at a scale of 1:25,000 (IGN 2018) to address the terrain accessibility of the wildfire site. These layers were subsequently grouped into three, each group depicting a level of accessibility, namely: (i) walking trails accessible by the ground crews only on foot, (ii) tracks encompassing roads for heavy trucks and are primarily unpaved, (iii) paved roads that permit movement of all types of vehicles and machinery. Furthermore, we added the fourth ground accessibility variable by obtaining the non-permanent and non-irrigated open croplands from the Corine Land Cover 2012 (European Environment Agency 2019). All metrics related to ground accessibility were constructed using the cost distance tool. This GIS-based tool calculates variation in distances as a function of a 3-dimensional distance space weighted by a cost layer. We used the terrain's angular slope, obtained from the Digital Elevation Model (IGN 2015) at a 30-m spatial resolution, as the cost function since all wildfire sites can be assessed on foot.

2.4.3. Fire spread obstacles

The cultural landscape of the Mediterranean is characterized by a diverse mix of forests, agricultural lands, open woodlands, grazing pastures, and anthropogenic features, forming a complex mosaic. This intricate composition gives rise to various clustering patterns, influencing firefighting strategies and shaping wildfire behaviour and spread rate. Despite the significant challenge posed by intense wildfire events that produce abundant embers, barriers and firebreaks remain essential for containing the spread of backing and flanking fires. Our study specifically examined the closeness of non-flammable obstacles, such as major roads, rivers, railways, and irrigation channels. Data concerning the types of barriers were collected from the BTN25 dataset (IGN 2018), which was integrated into a cohesive barrier network. Subsequently, we estimated the Euclidean distance to these barriers at a 30-meter resolution. Notably, we did not incorporate cost distance calculations in this analysis, as the spread rate and behaviour of wildfires are heavily influenced by factors such as terrain slope and the direction of wildfire propagation, information that was absent in the wildfire perimeter database.

2.4.4. Utilization of airborne support

Airborne support offers the flexibility to employ various firefighting strategies in remote regions, particularly where heavy machinery and equipment have limited ground access. Additionally, it is essential to directly combat heading wildfire intensities ranging from 1.2 to 2.4 meters in flame length, achievable solely through aerial intervention (Andrews et al. 2011). Suppression of wildfires by aerial intervention is attainable through direct water sprinkling on the Fireline or using chemicals like polyphosphate retardants, gels, and foams to create a temporary wildfire barrier (Plucinski and Pastor, 2013). We incorporated the influence of airborne support into our model by estimating the Euclidean distance of wildfire perimeter to heliports, air bases, and water reservoirs needed for efficient water supply (Silva et al. 2014). Vector layers containing information on airports, landing zones and water reservoir boundaries were obtained from BTN25 (IGN 2018).

2.4.5. Topographic contour

The topographical features of the terrain significantly shape wildfire behaviours. The configuration of ridges and valleys influences the implementation of firefighting strategies (Wei et al. 2019), with water divides often serving as critical boundaries that signal shifts in wildfire behaviour. Therefore, identifying these areas is strategically crucial for wildfire management (Otero et al. 2018). Wildfires progressing, for instance, on south-facing slopes often experience reduced intensity and spread rate upon crossing a water divide and entering north-facing terrain, where downhill wildfire spread has a limited impact on preheating fuel (Rothermel 1972). In our study, we incorporated the influence of terrain relief into our model to assess containment probabilities by integrating terrain curvature (Weiss 2001; Beier and Brost 2010). Terrain curvature, derived from a LiDAR-based digital elevation model with a resolution of 30 meters (IGN 2015), represents the terrain's shape based on slope. Curvature values near zero show flat areas such as plains, positive values show concave features (such as hills, ridges, and mountains), and negative values represent convex shapes (such as valleys, sinkholes, and basins).

2.5. Modelling containment

We utilized a presence/absence binomial classification approach to estimate the likelihood of wildfire containment. Specifically, we applied Random Forest, a non-parametric tree-based machine learning algorithm (Breiman 2001; Liaw and Wiener 2002). Random Forest functions as an ensemble of trees, each relying on a subset of random variables (Cutler et al. 2012), drawing upon the principles of binary recursive partitioning trees (Breiman 1984, 2001). This method employs sequential binary partitioning to divide the predictor space, generating descendant nodes until further partitioning is not feasible. While Random Forest is excellent at handling multi-class classification and regression tasks, it may be susceptible to overfitting. Mitigating this issue can be achieved through the implementation of bootstrapping and resampling techniques. Our model was trained, evaluated, and tested using the R environment's *caret* package (Kuhn 2008) and the *ranger* package (Wright and Ziegler 2017). To provide a basis for comparison and ascertain the selection of the best modelling algorithm, we trained additional models using logistic regression and support vector machine at *k-fold* ($k = 3$, 4, 5, and 6) cross-validation and tested the output against our Random Forest models.

2.5.1. Model calibration and training

We optimized our model for the best minimum observation in each node (*min node size*), and the number of predictors sampled at each split (*mtry*) by tuning the hyperparameters with a 10-fold cross-validation process repeated 5 times. The tunning was implemented by trying different combinations of *mtry* and min *node sizes* using the Gini index splitting rule and *Area Under the ROC* (receiver operating characteristic*) Curve* (AUC) as the metrics for validation, obtaining an *mtry* of 3 and a *min node size* of 5. The AUC is a threshold-independent metric that illustrates the performance of a binary classifier based on the comparison of the true positive rate against the false positive rate of a given model. Moreover, to avoid spatial autocorrelation of residuals and prevent model overfitting, the original sample dataset was split into training (80% of dataset) and testing (20% of dataset) using an additional *k-fold* ($k = 5$) cross-validation, resulting in 50 model realizations. Our workflow also included the calculation of residuals' spatial correlation through Moran's I index (Moran 1950), conducted over the 50 models to ensure the spatial conformity and independence of our results. All models, having satisfied the absence of spatial correlation, were combined into one final prediction corresponding to the median value of the predicted probability of the models. Additionally, we utilized the 'Leave One Out' method to understand the model's strength in capturing the temporal variation. In this method, we calibrated the model using all years but one and validated it using the year left out from the calibration. This process was repeated, leaving 1 year out at a time.

We initially tested the effectiveness of our existing calibration by considering it as a baseline model to address spatial autocorrelation. Our analysis revealed a stabilization range of approximately 1 km. To further remove spatial autocorrelation's influence and ensure its complete absence in our model, we constructed a training subset with a minimum distance set at 1 km.

2.5.2. Model performance evaluation

We assessed the performance of our model through *k-fold* cross-validation, as described earlier. To determine the predictive capability of each model, we utilized the test dataset and assessed the Area Under the Receiver Operating Characteristic Curve (AUC) (Hanley and McNeil 1982). The Receiver Operating Characteristic Curve (ROC) is a graphical representation that depicts the diagnostic accuracy of a binary classifier system across various discrimination thresholds. Similarly, the AUC measures the probability of a classifier assigning a higher rank to a randomly chosen positive instance compared to a randomly selected negative instance, with values ranging from 0.5 for random predictions to 1 for perfect predictions. Typically, models with an AUC exceeding 0.7 are deemed reliable (Zhou et al. 2011), while those surpassing 0.9 are considered excellent (McCune et al. 2002).

2.5.3. Variable performance evaluation

Machine learning algorithms are often termed as 'black boxes' due to their inherent characteristic of masking the process by which they generate predictions. Coincidentally, different methods have been developed to determine the role of covariates in model performance and the relationship between predictor and response variables. To explain how individual covariates contribute to model performance, we obtained each model's relative importance using node impurity, which measures the condition perfect for splitting tree branches during model training. The average of each predictor shows the measure of importance of the model.

Furthermore, we examined the relationship between each predictor and response variable using the Partial Dependency Plot (PDP) (Friedman 2001). This method provides a visual depiction of the incremental impact of a specific covariate on the predicted response. Using 2-D PDPs, we showed the marginal change in the model's response to a range of given values of the predictors using the Locally Weighted Scatterplot Smoothing function (LOESS) and represented the uncertainty of prediction using a 95% confidence interval. We built the PDPs using the *pdp* (Greenwell 2017) and *ggplot2* (Wickham 2016) R packages.

3. Results

3.1. Spatiotemporal pattern of containment success probability

We observed a gradient shift in containment success across the Catalonia landscape, with a decreasing probability towards the Lleida, Barcelona, Tarragona, and the Pyrenees Mountain range. The lower containment areas (≤ 0.75) are predominantly located in the north of Catalonia due to the presence of mountains, which limits accessibility and timely intervention of fire crews to contain wildfires. Similarly, mountainous peaks in central and southern Catalonia and the Pyrenees Mountain range in Girona exhibit a containment probability below 0.5, showing challenging areas of successful wildfire containment. Conversely, areas with high accessibility resulting from proximity to roads and predominantly were found to have high containment success, ranging between 0.75 to 1.0. At the centre of Catalonia, the aerial support's spatial impact manifested as a circular buffer, generated through linear distance computations to aircraft stations and water reservoirs, where the containment success was slightly higher within the radius of influence. Furthermore, containment probability also exhibits a temporal variation across the wildfire seasons in Catalonia, with a gradual decrease in mean probability from May to September (Figures 3 and 4).

3.2. Performance of the RF models

Following the AUC threshold classification by McCune et al. 2002, we obtained a good to excellent RF model, with all models having perpetually high AUC ranging from 0.75 to 0.91, with a mean AUC of 0.81 ± 0.03 1sd (Figure 5(*left*)). Additionally,

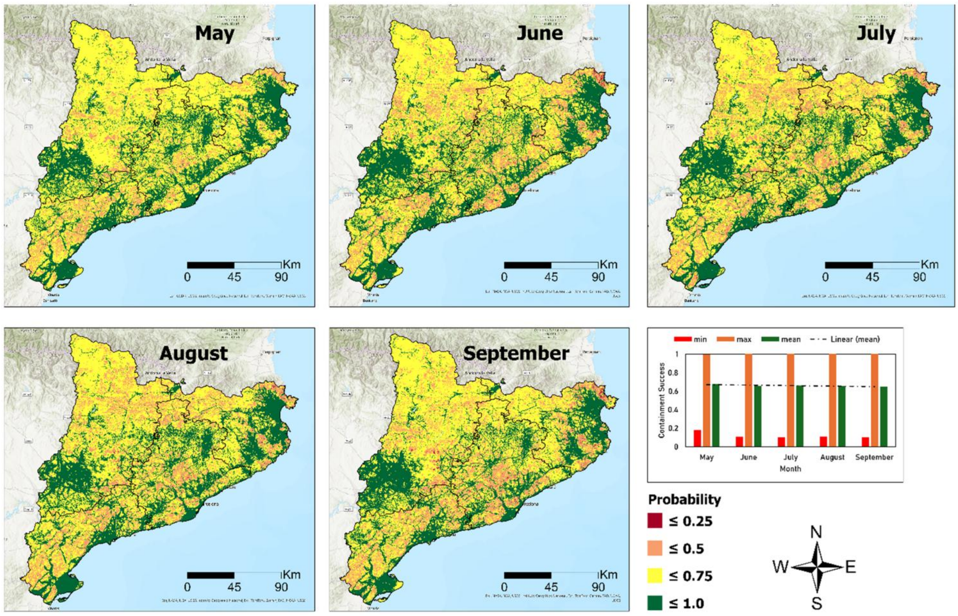

Figure 3. Spatial-temporal variation in containment probability across fire seasons in Catalonia. Each map represents the containment corresponding to each month and the black line represent the boundaries of each province. The bar chart presents the summary statistics of predicted probability in each month.

the model sufficiently captured the temporal variation across the years with a mean AUC of 0.72 ± 0.032 (Figure S9). Furthermore, we observed a clear distinction between the probability of containment success and failure, evident from the histogram density plot (Figure 5(*right*)), which has a consistently higher likelihood of containment success for presence locations. A larger frequency peak close to 0.70 was observed in the presence location following its gradual increase from 0.25, after which the curve gradually decreased without flattening out. In contrast, a density peak of 0.35 was obtained in the absence location, with the curve gradually tapering and flattening beyond the 0.5 probability threshold. Therefore, all parameters attest to the model's excellent predictive performance, reinforcing our findings' accuracy.

3.3. Performance of containment drivers

We assessed each predictor covariate's relative importance by ranking according to the percentage of relative node purity corresponding to each variable (Figures 6 and 7). The cost distance to roads exerted the most significant influence and was consistently 100% important, although some outliers showed less than 100% importance. Temperature and windspeed are closely followed, with a median relative influence of around 85% and 65%, respectively. Distance to airborne support followed closely to complete the upper position of relative influence. In the mid-tier position are distance to crops, cost distance to tracks and distance to watershed, yielding median relative importance close to 45%. Distance to barriers, northness, cost distance to paths and curvature shows the lowest influence on the model's predictive capacity, with a

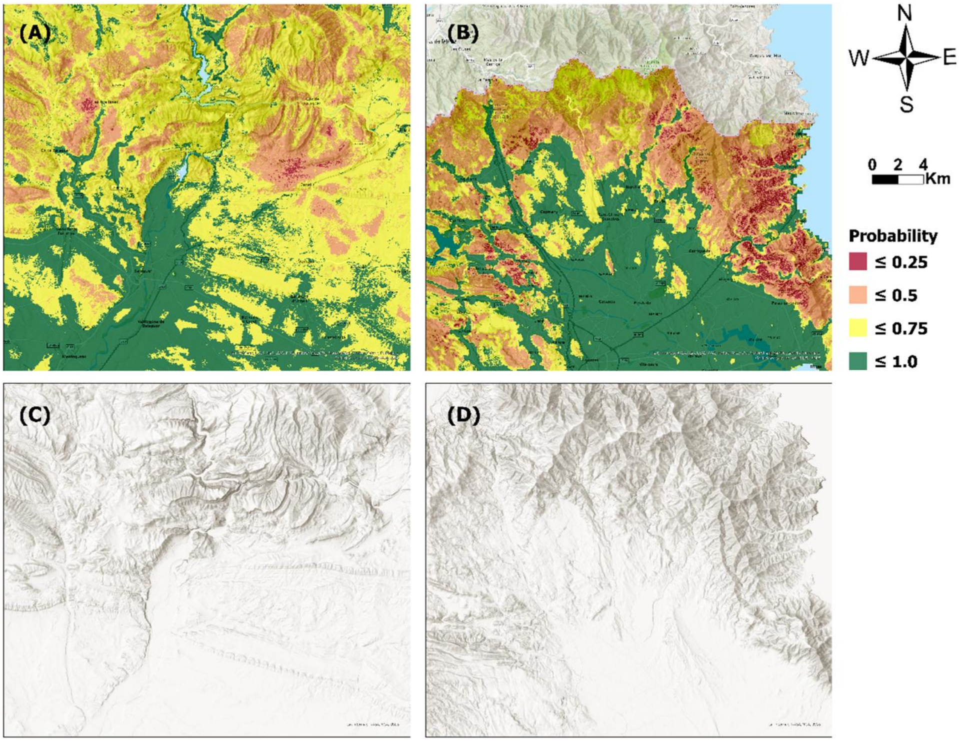

Figure 4. A close-up view (1:100,000) of the variation in containment probability between the plains and mountainous part of (A) Lleida, (B) Girona. A hill shade was used as a backdrop to show topographical variation; (C) and (D) represent the hill shade backdrop of a and B, respectively.

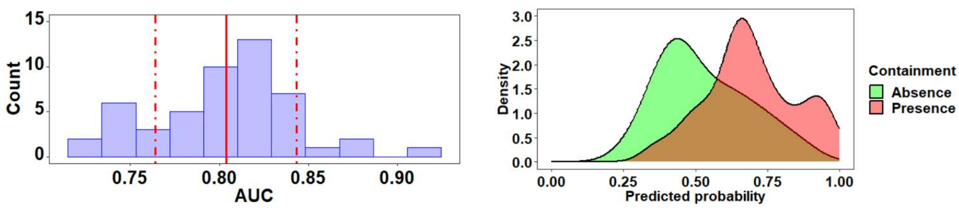

Figure 5. Frequency distribution of AUC values from RF models (*left*) and density histogram of predicted probability for escaped (green) and contained (red) fires. The solid red line indicates the mean AUC and dashed red lines indicate ± 1 standard deviation.

median importance less than 25%. Isolating the weather variables, we observed temperature as the most influential weather variable, followed by wind speed.

We further analyzed and summarized each predictor covariate's marginal effect in predicting the containment probability of wildfires that escaped the initial attack in Catalonia using separate partial dependency plots (Figure 8). From our results, we isolated critical traits of areas with a higher containment probability to include those having (i) proximity to the roadside with (ii) low to moderate daily temperature and, (iii) low level of windspeed, (iv) within a 2 km distance to Heliport, and close to (v) agricultural fields and (vi) tracks for heavy truck, (vii) proximity to watershed, and

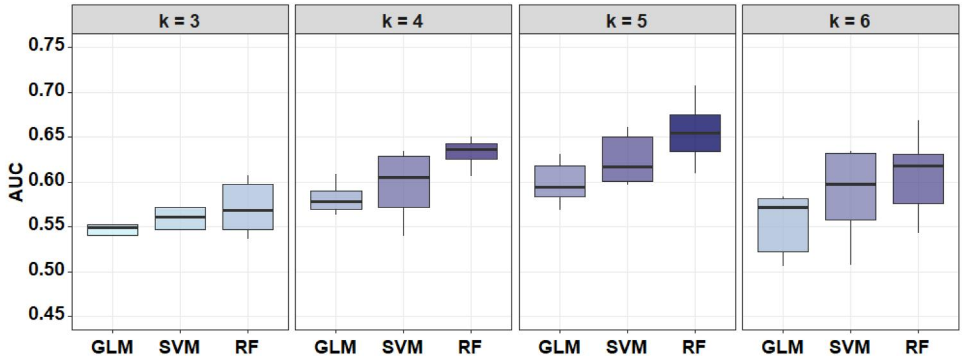

Figure 6. Comparative summary of model performance across different *k-fold* data resampling of each modelling algorithms. The boxplot represents the 1st and 3rd quartiles, Middle line shows the median value, colour alpha represents median variation within each *k-fold*, the whiskers represent 10th and 90th percentile. GLM = generalised linear model; SVM = support vector machine; RF = random Forest.

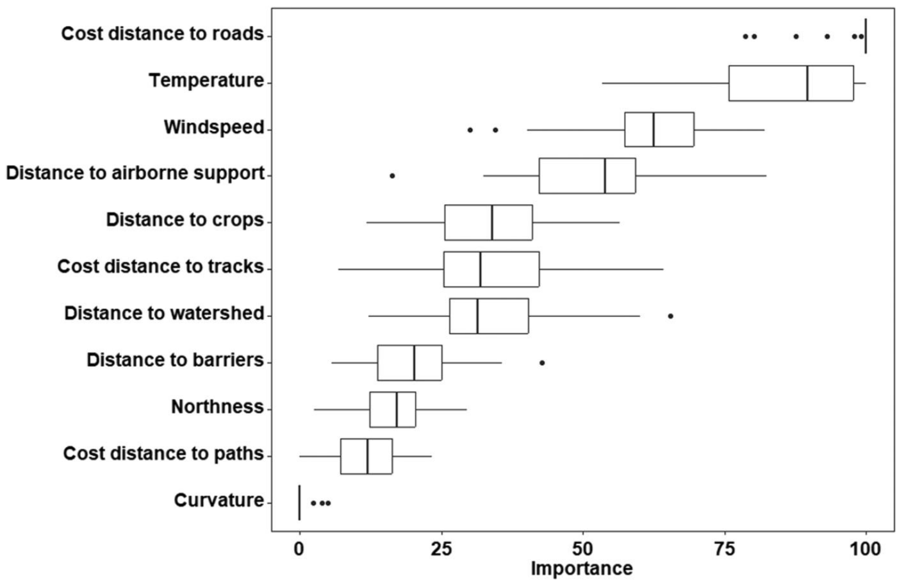

Figure 7. Boxplot distribution of the variable performance using node purity. The boxplot represents the 1st and 3rd quartiles, Middle line shows the median value, the whiskers represent 10th and 90th percentile, the dots represent outliers (values below 10th percentile or above 90th percentile).

(viii) to a fire spread barrier while, (ix) far from walking trails. Considering the terrain undulation, we observed a high containment probability in areas around plains and negative curvatures, such as valleys, basins, or sinkholes.

4. Discussion

Wildfires exhibit extreme behaviour with severe destructive power due to the increasing duration of hot and dry seasons (Pastor et al. 2020; Moreno et al. 2023).

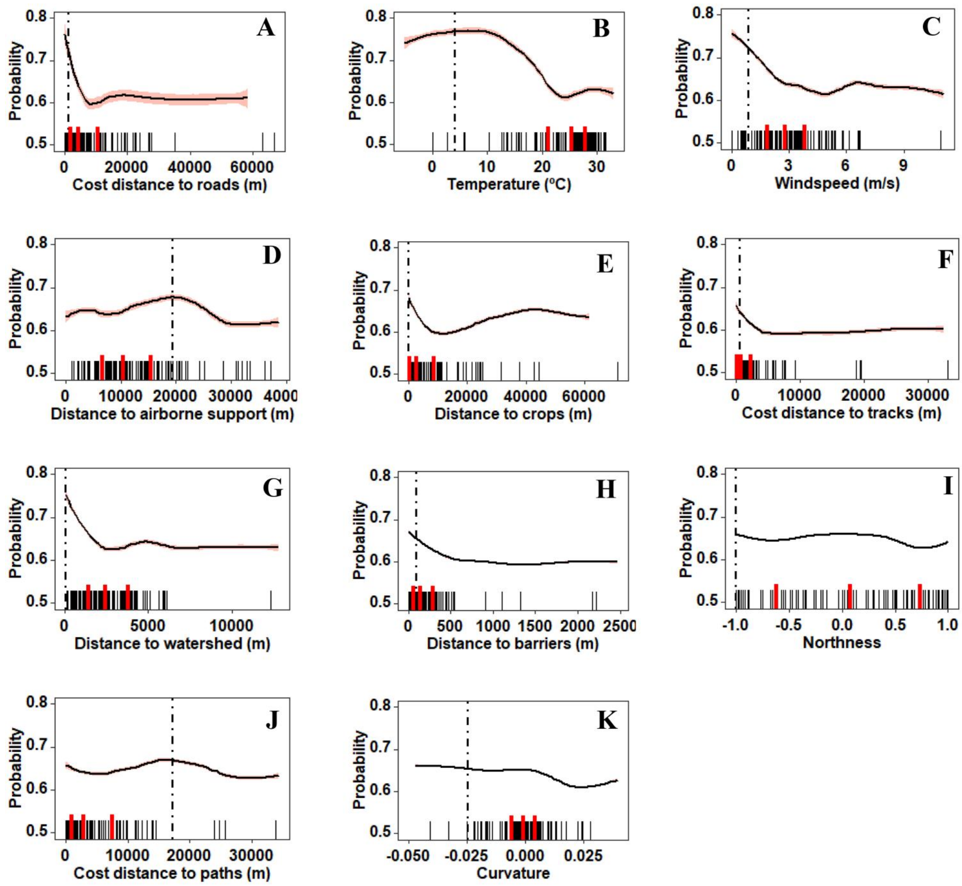

Figure 8. Partial dependence plots of (A) cost distance to roads, (B) Temperature, (C) Windspeed, (D) Distance to airborne support, (E) Distance to crop, (F) Distance to tracks, (G) Distance to watershed, (H) Distance to barriers, (I) Northness, and (J) Cost distance to paths, and (K) Curvature. The x-axis denotes the covariate values, and the y-axis shows the corresponding predicted probability. The vertical dashed lines show the point of highest containment probability. The solid line represents the averaged dependence whereas shaded areas represent the standard error. The subset of the density of presence events is represented by vertical black lines above the x-axes, and overlaid red ticks depict the quartiles.

Firefighters are often responsible for reacting simultaneously to wildfire suppression, structural protection, and evacuation during a wildfire outbreak. Combining these can overwhelm the firefighter capacity, increasing casualties, especially when suppression is unsuccessful, and wildfire escapes rapidly (Pastor et al. 2020). Nevertheless, firefighting is crucial in preventing property loss and reducing burned areas in extreme wildfires (Penman et al. 2015; van der Merwe et al. 2015). Ensuring a safe and effective response to wildfires has emerged as a critical concern for minimizing damages and fostering harmonious coexistence with wildfire in southern Mediterranean regions and other wildfire-prone areas (van der Merwe et al. 2015; Alcasena et al. 2019). While initial attack strategies are pivotal in identifying and suppressing unplanned wildfire ignitions that may escape, decision-making regarding resource allocation and operational planning often relies on expert-driven criteria developed from historical data, introducing inherent uncertainty due to the multitude of complex and evolving factors influencing successful responses (Rodrigues et al. 2019). By identifying areas favourable for containment and effectively deploying resources, a proactive response can be undertaken by wildfire managers (Syphard et al. 2014; Gonzalez-Olabarria et al. 2019). Still, the limited response time and the cascading effect of weather and biophysical variables makes efficient deployment of equipment and personnel difficult during the period of active wildfires. Therefore, there is the need for scientific tools to improve firefighters and managers understanding of the spatiotemporal chances of containment success during the period of active wildfires. In view of this, we developed a model of wildfire containment probability in Catalonia using the historical wildfire perimeters and detailed biophysical and weather variables associated with wildfire spread and control. We presented the results as fine-resolution raster layers to facilitate improve containment effort for wildfires that escape initial attack in Catalonia.

Models addressing initial attack and containment have been documented in the literature (Finney et al. 2009; Rodrigues et al. 2019, 2020b), although few studies have developed quantitative and empirical wildfire containment models in the Mediterranean (Wei et al. 2019). Our methodological workflow aligns with the Containment modelling of Rodrigues et al. 2020b in Catalonia, relying on calibrating a binary model of containment probability from a set of detailed geospatial explanatory variables using statistical and machine learning algorithms. Our model performed similarly to this previous model and mostly agrees with the drivers' behaviour and wildfire containment's spatial distribution.

The novelty in our research, thereby addressing the limitation of the previous study, is incorporating daily temperature and windspeed and spatializing the output at a satisfactory resolution of 30 m. This allowed us to capture the meteorological effect and develop a method to allocate daily data based on the spatial extent of the wildfire perimeter and burn duration. Additionally, by calibrating our model with daily weather data and validating with independent fire seasons in the temporal cross-validation to evaluate its operational implementation while capturing climate variability, our result presented a temporal dimension to underscore the crucial role of seasonal changes in wildfire containment success in Catalonia.

We identified the cost distance to roads as the most influential factor in successfully containing wildfires that escape initial attacks in Catalonia, emphasizing the importance of surface accessibility, primarily through roads that permit the movement of all vehicle types. Usually, the primary suppression effort is generally executed using land-based equipment, whose transportation to the wildfire site is done through roads. Therefore, an adequate road network ensures the timely arrival of equipment and machinery at the wildfire site, promoting successful containment. Conversely, a distant or remote wildfire location with inadequate road accessibility may significantly prolong response time, leading to uncontrolled wildfire growth and escalation. We observed this trend in the spatial-temporal pattern of containment probability in Catalonia, where success probability was significantly lower in remote areas dominated primarily by mountains, especially in the Pyrenees Mountain range. In contrast, regions exhibiting dense road networks with proximity to agricultural plains and urbanized areas, such as central Lleida, the agricultural plain of Girona and the cosmopolitan area of Barcelona, have a high containment probability due to increased accessibility and a characteristically limited fuel load (Figure 4 and Figure S4).

All weather variables highlight the significant impact of extreme weather conditions on containment success and the opportunistic nature of wildfire propagation. Hazardous meteorological factors such as extreme temperatures, drought, and dense vegetation have been linked to the rapid escalation of large wildfires (Moreno et al. 2023), consequently diminishing containment success rates. Additionally, periods characterized by extreme weather events, marked by elevated temperatures and wind speeds, can hamper firefighting efforts, thereby increasing the likelihood of wildfires evading initial containment measures (Finney et al. 2009; Rodrigues et al. 2019). Through the isolation of individual variables, as observed from the partial dependency plots (Figure 8(B,C)) of our modelling output, we have identified optimal conditions for achieving containment success: temperatures within 10 °C and wind speeds below $3\,\mathrm{ms}^{-1}$. Extreme temperatures, notably when combined with reduced precipitation, can drastically decrease fuel moisture content, resulting in heightened fuel loads that are challenging to manage upon ignition. The windy condition often coincides with low relative humidity, facilitating wildfire spread through flame fanning and ember transport. Conversely, when a windy condition coincides with a high relative humidity, it can impede water-based firefighting efforts by reducing the water absorption capacity of saturated fuels and posing challenges for firefighters attempting to penetrate the fire core for effective extinguishment while wildfire spreads rapidly. Moreover, high humidity levels typically correlate with reduced visibility due to the formation of thick smoke and fog, hindering accurate targeting of water streams onto the fires.

We demonstrated the superior performance of random forest over logistic regression and support vector machine in probability modelling (Figure 6), reinforcing past research findings (Rodrigues et al. 2019, 2020b; Gelabert et al. 2022). We limited the window of our comparison to the traditional logistic regression and machine learning techniques without exploring the deep learning algorithm approach. This decision is based on past studies documenting RF's excellent performance, robustness, and suitability in wildfire science (Sakr et al. 2011; Bar Massada et al. 2012; Amatulli et al. 2013; Sanabria et al. 2013; Rodrigues and de la Riva 2014). While we achieve substantially high accuracy in our models, further studies can benefit from comparing our results with other deep learning algorithms and exploring the variation in prediction capacity. Nevertheless, our result has provided substantial information that benefits fire managers, response crews, and policymakers regarding effective wildfire response and overall wildfire management.

The current exclusion policy, which mandates the suppression of all fires, can be challenging to maintain due to the continuous need for financial resources and budget allocations which cannot be sustained indefinitely (Rodrigues et al. 2019). Moreover, its efficacy in mitigating large wildfire occurrences has proven minimal (San-Miguel-Ayanz et al. 2013). Besides, several factors such as wildfire-favoured weather conditions, wildfire detection time, arrival time and productivity of fire crew, and wildfire spread rate can limit suppression effort at the initial wildfire stage (Gonzalez-Olabarria et al. 2019), thereby increasing the likelihood of a wildfire escalating beyond control. Furthermore, there is a growing acknowledgement that aggressively suppressing all small wildfires may have inadvertently exacerbated the issue of fuel accumulation in Mediterranean landscapes over recent decades (Otero and Nielsen 2017), thereby shifting the wildfire regime from fuel-limited to flammability-limited wildfire regime. Historically, vegetation in the Mediterranean has evolved with wildfire and is highly wildfire-adapted (Viana-Soto et al. 2017) due to its ecological characteristics, including the capacity to create a persistent seed bank that opens after a wildfire event (Pérez-Cabello et al. 2009). However, a shift in the historical wildfire regime will increase wildfire severity due to wildfire-favoured weather conditions, thus affecting ecological composition, plant species distribution and alterations in ecosystem functions and forest structure (Pausas and Fernández-Muñoz 2012). Considering these underlying circumstances, preventing large-scale wildfires can be practically challenging. In the Mediterranean, wildfire response is critical, especially during the height of wildfire seasons when weather conditions combined with landscape complexity promote large wildfires. In this regard, our probability maps provide valuable and practical information towards promoting efficient response and protecting against economic loss and environmental damage. However, we note a caution in the practical application of our maps on the local scale; areas showing a high containment level can fall short of the safety requirements for firefighters (Campbell et al. 2019; Page and Butler 2019) by not adequately providing escape routes since factors relating to safety were excluded from the model. For example, areas with a high containment probability due to proximity to the roadside and plain terrain might not have adequate safety space for fire crew during suppressing efforts. Therefore, understanding the local conditions within the context of crew safety is essential for effectively using our maps.

Our wildfire containment model serves as a valuable tool for guiding proactive firefighting efforts. For example, it allows us to assess and quantify potential improvements in predicted wildfire containment probabilities across various scenarios. This includes evaluating the effectiveness of new firefighting infrastructure projects, such as developing forest track networks, establishing water pond systems, and implementing fuel management plans to enhance wildfire suppression capabilities in strategic areas. Moreover, considering the substantial investment in maintaining existing infrastructures, it may be practical to explore more cost-effective alternatives. This could involve replacing ageing infrastructure with newer, more efficient solutions, thus optimizing resource allocation for wildfire containment efforts. Besides, our containment maps can guide the establishment of strategic fire control zones to enhance operational decisions and improve firefighting efficiency. During concurrent wildfire events, firefighters can strategically allocate firefighting resources based on the probability of containing the fires, thereby ensuring optimal use of resources with desirable outcomes.

While our findings largely align with those of Rodrigues et al. (2020b), we observed variations in the behaviour of specific drivers. These deviations may stem from enhancements made to our model, which now incorporates temporal weather fluctuations and utilizes a different methodological workflow and variable selection process. However, it is worth noting that these differences have minimal impact on the overall outcomes, as both models yield spatially comparable results. Moreover, our model represents a significant advancement over prior approaches. By addressing the critical limitations of previous models, ours introduces a novel methodology that provides a nuanced understanding of containment probability, particularly within the context of the wildfire season in Catalonia. This enhanced method enables us to capture both temporal and spatial variations in a manner that enhances the robustness and applicability of our findings.

As previously discussed, we opted not to include variables related to crew safety in our model to maintain its clarity and robustness. Also, safety zone designs necessitate different factors relating to heat transfer and smoke (Butler 2014; Page and Butler 2017); capturing them further complicates the model. However, future research could explore these safety-related variables in a spatial context, allowing for a more comprehensive understanding of their impact. Comparing these findings with our results could provide valuable insights. Besides, although our methodology captured the variation in weather conditions across the fire-burning days, it assumes a constant wildfire spread rate. This can be a potential limitation in the spatial representation of containment probability, as wildfires do not necessarily spread at a fixed rate. Overcoming this requires real-time monitoring of weather variations during wildfires. This can be done, for example, by extracting real-time weather data from providers' websites and implementing a periodical model retraining using historical and real-time data. Further studies could also explore models that accurately simulate weather changes during wildfire events. Finally, integrating our model with existing wildfire simulation models presents an opportunity to elucidate landscape susceptibility to wildfires, how environmental factors contribute to wildfire growth and the chance of effectively containing these wildfires. This holistic approach can enhance preparedness efforts by providing insights into potential wildfire scenarios and facilitating more effective response strategies. Ultimately, this collaborative approach aims to safeguard lives, properties, and critical ecosystems from the devastating effects of wildfires.

5. Conclusion

Amidst the escalating wildfire crisis in the Mediterranean, effective management is inevitable. The failure of the fire exclusion policy and the resulting accumulation of fuel due to the complete suppression of minor wildfires have left the landscape susceptible to increasingly severe mega-wildfire events. While comprehensive wildfire prevention efforts, leveraging the expertise of local communities, firefighters, and sufficient resources, are essential, the unique landscape characteristics of the Mediterranean render total wildfire prevention unattainable. Instead, a more feasible approach entails focusing on wildfire suppression, mainly targeting large wildfires, which account for most burned areas in the Mediterranean region. In this regard, we developed a study on the containment probability of escaped wildfires in Catalonia, a region characterized by a Mediterranean climate. By addressing the limitations of an existing model, we have enhanced our methodology to incorporate daily weather conditions and produce results at a fine-resolution scale. Our RF model has demonstrated superior performance over other classification and regression models, such as logistic regression and support vector machine, across various data resampling techniques. The high accuracy of our RF-based model underscores its efficiency in predicting containment probabilities. The containment maps generated at a 30-meter resolution for multiple wildfire seasons in Catalonia offer valuable insights for guiding management operations and decision-making processes. However, it is essential to note that our maps do not account for variables related to firefighter safety. Therefore, their usage should be accompanied by a comprehensive understanding of Catalonia's landscape. Furthermore, integrating our findings with wildfire simulation models present numerous opportunities for advancing research in understanding wildfire dynamics across landscapes and identifying strategic containment areas. Our framework and workflow can serve as a blueprint for developing similar containment models tailored to different regions, climatic conditions, ecosystem configurations, and landscape characteristics.

Acknowledgement

The funding for the research was provided by the European Union's scholarship support through the Erasmus Mundus Joint Master's Degree program.

Authors' contributions

Alawode, Lawrence Gbenga: Conceptualization, Methodology, Data curation, Investigation, Formal analysis, Software, Visualization, Writing- Original draft preparation, Validation; Gelabert Vadillo, Joan Pere: Conceptualization, Methodology, Data curation, Software, Validation, Reviewing, Editing, Supervision; Rodrigues, Marcos: Conceptualization, Methodology, Investigation, Supervision, Reviewing and Editing. All authors reviewed the manuscript.

Disclosure statement

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

Funding

This work was financed by the projects FIREPATHS (PID2020-116556RA-I00), funded by the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033); and FireCycle (CNS2023-144228), funded by the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033) and the European Union ('NextGenerationEU'/PRTR).

Data availability statement

All datasets used and analysed during this study are available in the article and from the corresponding author on request.

References

Alcasena FJ, Ager AA, Bailey JD, Pineda N, Vega-García C. 2019. Towards a comprehensive wildfire management strategy for mediterranean areas: framework development and implementation in Catalonia, Spain. J Environ Manage. 231:303–320. doi: 10.1016/j.jenvman.2018.10.027.

Alcasena FJ, Evers CR, Vega-Garcia C. 2018. The wildland-urban interface raster dataset of Catalonia. Data Brief. 17:124–128. doi: 10.1016/j.dib.2017.12.06.

Amatulli G, Camia A, San-Miguel-Ayanz J. 2013. Estimating future burned areas under changing climate in the EU-mediterranean countries. Sci Total Environ. 450:209–222. doi: 10.1016/j.scitotenv.2013.02.014.

AMB. 2023. Àrea metropolitana de Barcelona [WWW document]. Población -Àrea Metrop, Barcelona. [accessed 2023 November 29]. http://www.amb.cat/s/web/area-metropolitana/coneixer-l-are a-metropolitana/poblacio.html

Andrews PL, Heinsch FA, Schelvan L. 2011. How to generate and interpret fire characteristics charts for surface and crown fire behavior. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Arienti MC, Cumming SG, Boutin S. 2006. Empirical models of forest fire initial attack success probabilities: the effects of fuels, anthropogenic linear features, fire weather, and management. Can J For Res. 36(12):3155–3166. doi: 10.1139/x06-188.

Badia A, Pallares-Barbera M, Valldeperas N, Gisbert M. 2019. Wildfires in the wildland-urban interface in Catalonia: vulnerability analysis based on land use and land cover change. Sci Total Environ. 673:184–196. doi: 10.1016/j.scitotenv.2019.04.012.

Bar Massada A, Syphard AD, Stewart SI, Radeloff VC. 2012. Wildfire ignition distribution modelling: a comparative study in the Huron–Manistee National Forest, Michigan, USA. Int. J. Wildl. Fire. 22(2):174–183. doi: 10.1071/WF11178.

Beier P, Brost B. 2010. Use of land facets to plan for climate change: conserving the arenas, not the actors. Conserv Biol. 24(3):701–710. doi: 10.1111/j.1523-1739.2009.01422.x.

Breiman L. 1984. Classification and regression trees. London (UK): Champan & Hall.

Breiman L. 2001. Random forests. Machine Learning. 45(1):5–32. doi: 10.1201/9780429469275-8.

Butler BW. 2014. Wildland firefighter safety zones: a review of past science and summary of future needs. Int J Wildland Fire. 23(3):295–308. doi: 10.1071/WF13021.

Calheiros T, Pereira MG, Nunes JP. 2021. Assessing impacts of future climate change on extreme fire weather and pyro-regions in Iberian Peninsula. Sci Total Environ. 754:142233. doi: 10.1016/j.scitotenv.2020.142233.

Campbell MJ, Page WG, Dennison PE, Butler BW. 2019. Escape route index: a spatially-explicit measure of wildland firefighter egress capacity. Fire. 2(3):40. doi: 10.3390/fire2030040.

Castellnou M, Pagés J, Miralles M, Piqué M. 2009. Tipificación de los incendios forestales de Cataluña. Elaboración del mapa de incendios de diseño como herramienta para la gestión forestal. In: 50 Congreso Forestal, Ávila; p. 15.

Cos J, Doblas-Reyes F, Jury M, Marcos R, Bretonnière P-A, Samsó M. 2022. The mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. Earth Syst Dyn. 13(1):321–340. doi: 10.5194/esd-13-321-2022.

Cutler A, Cutler DR, Stevens JR. 2012. Ensemble machine learning. New York (NY): Springer. doi: 10.1007/978-1-4419-9326-7.

Díaz-Avalos C, Juan P, Serra-Saurina L. 2016. Modeling fire size of wildfires in Castellon (Spain), using spatiotemporal marked point processes. For Ecol Manage. 381:360–369. doi: 10.1016/j.foreco.2016.09.013.

Díaz-Delgado R, Lloret F, Pons X. 2004a. Spatial patterns of fire occurrence in Catalonia, NE, Spain. Landscape Ecol. 19(7):731–745. doi: 10.1007/s10980-005-0183-1.

Díaz-Delgado R, Lloret F, Pons X. 2004b. Statistical analysis of fire frequency models for Catalonia (NE Spain, 1975–1998) based on fire scar maps from Landsat MSS data. Int J Wildland Fire. 13(1):89–99. doi: 10.1071/WF02051.

EGIF. 2024. Estadística General de Incendios Forestales. Ministerio para la Transición Ecológica y el Reto Demográfico Plaza San Juan de la Cruz, 10 28071 Madrid. https://www.miteco.gob.es/en/biodiversidad/temas/inventarios-nacionales/inventario-espanol-patrimonio-natural-biodiv/sistema-indicadores/06c-estadistica-incendios-forestales.html.
European Commission: Directorate-General for Environment, Joint Research Centre, Libertà G, San-Miguel-Ayanz J, Durrant T, et al. 2016. Forest fires in Europe, Middle East and North Africa 2015, Publications Office, 2016. https://data.europa.eu/doi/10.2788/914.
European Environment Agency. 2019. Corine land cover (CLC) 2012. Version 20b [Internet]. [cited 2018 Dec 22]. https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012.
Fernandes PM, Pacheco AP, Almeida R, Claro J. 2016. The role of fire-suppression force in limiting the spread of extremely large forest fires in Portugal. Eur J Forest Res. 135(2):253–262. doi: 10.1007/s10342-015-0933-8.
Finney M, Grenfell IC, McHugh CW. 2009. Modeling containment of large wildfires using generalised linear mixed-model analysis. For Sci. 55(3):249–255. doi: 10.1093/forestscience/55.3.249.
Friedman JH. 2001. Greedy function approximation: a gradient boosting machine. Ann Statist. 29(5):44. doi: 10.1214/aos/1013203451.
Galizia LF, Barbero R, Rodrigues M, Ruffault J, Pimont F, Curt T. 2023. Global warming reshapes European pyroregions. Earth's Future. 11(5):e2022EF003182. doi: 10.1029/2022EF003182.
Gelabert PJ, Rodrigues M, Vidal-Macua JJ, Ameztegui A, Vega-Garcia C. 2022. Spatially explicit modeling of the probability of land abandonment in the Spanish Pyrenees. Landscape Urban Plann. 226:104487. doi: 10.1016/j.landurbplan.2022.104487.
GENCAT. 2024. Base cartográfica de les zones cremades per incendis forestals produïts entre el 1986 i el 2021. Departament d'Agricultura, Ramaderia, Pesca i Alimentacío. Generalitat de Catalunya. https://agricultura.gencat.cat/ca/serveis/cartografia-sig/bases-cartografiques/boscos/incendis-forestals/.
Gill AM, Stephens SL, Cary GJ. 2013. The worldwide "wildfire" problem. Ecol Appl. 23(2):438–454. doi: 10.1890/10-2213.1.
Gonçalves AC, Sousa AMO. 2017. The fire in the mediterranean region: a case study of forest fires in Portugal. In Fuerst-Bjelis B, editor. Mediterranean identities - environment, society, culture [Internet]. InTech; [accessed 2024 Dec 27]. doi: 10.5772/intechopen.69410.
González JR, Pukkala T. 2007. Characterisation of forest fires in Catalonia (north-east Spain). Eur J Forest Res. 126(3):421–429. doi: 10.1007/s10342-006-0164-0.
Gonzalez-Olabarria JR, Brotons L, Gritten D, Tudela A, Teres JA. 2012. Identifying location and causality of fire ignition hotspots in a mediterranean region. Int J Wildland Fire. 21(7):905–914. doi: 10.1071/WF11039.
Gonzalez-Olabarria JR, Mola-Yudego B, Coll L. 2015. Different factors for different causes: analysis of the spatial aggregations of fire ignitions in Catalonia (Spain). Risk Anal. 35(7):1197–1209. doi: 10.1111/risa.12339.
Gonzalez-Olabarria JR, Reynolds KM, Larrañaga A, Garcia-Gonzalo J, Busquets E, Pique M. 2019. Strategic and tactical planning to improve suppression efforts against large forest fires in the Catalonia region of Spain. For Ecol Manage. 432:612–622. doi: 10.1016/j.foreco.2018.09.039.
Greenwell BM. 2017. pdp: an R package for constructing partial dependence plots. R J. 9(1):421–436. https://journal.r-project.org/archive/2017/RJ-2017-016/index.html. doi: 10.32614/RJ-2017-016.
Handmer J, Proudley B. 2008. The economics of interface wildfires. USDA Forest Service, Pacific Southwest Research Station, General Technical Report PSWGTR-208. Albany: USDA; p. 627–637.
Hanley JA, McNeil BJ. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 143(1):29–36. doi: 10.1148/radiology.143.1.7063747.
IDESCAT. 2023. Institut d'Estadística de Catalunya [WWW Document]. [accessed 2023 November 29]. https://www.idescat.cat

Kuhn M. 2008. Building predictive models in R using the caret package. J Stat Software. 28(5):1–26. doi: 10.18637/jss.v028.i05.
Liaw A, Wiener M. 2002. Classification and regression by randomForest. R News. 2:18–22.
Lopez-Bustins JA, Arbiol-Roca L, Martin-Vide J, Barrera-Escoda A, Prohom M. 2020. Intra-annual variability of the Western mediterranean oscillation (WeMO) and occurrence of extreme torrential precipitation in Catalonia (NE Iberia). Nat Hazards Earth Syst Sci. 20(9):2483–2501. doi: 10.5194/nhess-20-2483-2020.
MAPAMA. n.d. Estadística General de Incendios Forestales. http://www.mapama.gob.es/es/desarrollorural/estadisticas/Incendios_default.aspx.
Martínez J, Vega-Garcia C, Chuvieco E. 2009. Human-caused wildfire risk rating for prevention planning in Spain. J Environ Manage. 90(2):1241–1252. doi: 10.1016/j.jenvman.2008.07.005.
McCune B, Grace JB, Urban DL. 2002. Analysis of ecological communities. Glenden Beach (OR): MJM Software Design.
Moran PAP. 1950. Notes on continuous stochastic phenomena. Biometrika. 37(1–2):17–23. doi: 10.1093/biomet/37.1-2.17.
Moreno M, Bertolín C, Arlanzón D, Ortiz P, Ortiz R. 2023. Climate change, large fires, and cultural landscapes in the mediterranean basin: an analysis in southern Spain. Heliyon. 9(6):e16941. doi: 10.1016/j.heliyon.2023.e16941.
Muñoz Sabater J. 2019. ERA5-land hourly data from 2001 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi: 10.24381/CDS.E2161BAC.
O'Connor CD, Calkin DE, Thompson MP. 2017. An empirical machine learning method for predicting potential fire control locations for pre-fire planning and operational fire management. Int J Wildland Fire. 26(7):587–597. doi: 10.1071/WF16135.
O'Hara KL. 2014. Multiaged silviculture: managing for complex forest stand structures. Oxford: Oxford University Press; p. 213. doi: 10.1093/acprof:oso/9780198703068.001.0001.
Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A, Pereira, JMC, 2. 2012. 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.
Otero I, Castellnou M, González I, Arilla E, Castell L, Castellví J, Sánchez F, Nielsen JØ. 2018. Democratizing wildfire strategies. Do you realize what it means? Insights from a participatory process in the Montseny region (Catalonia, Spain). PLoS One. 13(10):e0204806. doi: 10.1371/journal.pone.0204806.
Otero I, Nielsen JØ. 2017. Coexisting with wildfire? Achievements and challenges for a radical social-ecological transformation in Catalonia (Spain). Geoforum. 85:234–246. doi: 10.1016/j.geoforum.2017.07.020.
Page WG, Butler BW. 2017. An empirically based approach to defining wildland firefighter safety and survival zone separation distances. Int J Wildland Fire. 26(8):655–667. doi: 10.1071/WF16213.
Pastor E, Muñoz JA, Caballero D, Àgueda A, Dalmau F, Planas E. 2020. Wildland–urban interface fires in Spain: summary of the policy framework and recommendations for improvement. Fire Technol. 56(5):1831–1851. doi: 10.1007/s10694-019-00883-z.
Pausas JG, Fernández-Muñoz S. 2012. Fire regime changes in the Western mediterranean basin: from fuel-limited to drought-driven fire regime. Clim Change. 110(1-2):215–226. doi: 10.1007/s10584-011-0060-6.
Penman TD, Nicholson AE, Bradstock RA, Collins L, Penman SH, Price OF. 2015. Reducing the risk of house loss due to wildfires. Environ Modell Software. 67:12–25. doi: 10.1016/j.envsoft.2014.12.020.
Pereira MG, Trigo RM, da Camara CC, Pereira JM, Leite SM. 2005. Synoptic patterns associated with large summer forest fires in Portugal. Agric For Meteorol. 129(1-2):11–25. doi: 10.1016/j.agrformet.2004.12.007.
Pérez-Cabello F, Echevarría MT, Ibarra P, De la Riva J. 2009. Effects of fire on vegetation, soil and hydrogeo-morphological behavior in mediterranean ecosystems. In: Chuvieco E (editor).

Earth observation of wildland fires in mediterranean ecosystems. Berlin/Heidelberg (Germany): Springer; p. 111–128. doi: 10.1007/978-3-642-01754-4.
Petrovic N, Carlson JM. 2012. A decision-making framework for wildfire suppression. Int J Wildland Fire. 21(8):927–937. doi: 10.1071/WF11140.
Pineda N, Rigo T. 2017. The rainfall factor in lightning ignited wildfires in Catalonia. Agric For Meteorol. 239:249–263. doi: 10.1016/j.agrformet.2017.03.016.
Plucinski MP. 2012. Factors affecting containment area and time of Australian forest fires featuring aerial suppression. For Sci. 58(4):390–398. doi: 10.5849/forsci.10-096.
Plucinski MP, Pastor E. 2013. Criteria and methodology for evaluating aerial wildfire suppression. Int J Wildland Fire. 22(8):1144. doi: 10.1071/WF13040.
Riley KL, Thompson MP, Scott JH, Gilbertson-Day JW. 2018. A model-based framework to evaluate alternative wildfire suppression strategies. Resources. 7(1):4. doi: 10.3390/resources7010004.
Rodrigues M, Alcasena F, Vega-García C. 2019. Modeling initial attack success of wildfire suppression in Catalonia, Spain. Sci Total Environ. 666:915–927. doi: 10.1016/j.scitotenv.2019.02.323.
Rodrigues M, Alcasena F, Gelabert P, Vega-García C. 2020b. Geospatial modeling of containment probability for escaped wildfires in a mediterranean region. Risk Anal. 40(9):1762–1779. doi: 10.1111/risa.13524.
Rodrigues M, de la Riva J. 2014. An insight into machine-learning algorithms to model human-caused wildfire occurrence. Environ Model Software. 57:192–201. doi: 10.1016/j.envsoft.2014.03.003.
Rodrigues M, Jiménez-Ruano A, de la Riva J. 2020a. Fire regime dynamics in mainland Spain. Part 1: drivers of change. Sci Total Environ. 721:135841. doi: 10.1016/j.scitotenv.2019.135841.
Rothermel RC. 1972. A mathematical model for predicting fire spread in wildland fuels. Department of Agriculture, Intermountain Forest and Range.
Sakr GE, Elhajj IH, Mitri G. 2011. Efficient forest fire occurrence prediction for developing countries using two weather parameters. Eng Appl Artif Intell. 24(5):888–894. doi: 10.1016/j.engappai.2011.02.017.
Sanabria LA, Qin X, Li J, Cechet RP, Lucas C. 2013. Spatial interpolation of McArthur's forest fire danger index across Australia: observational study. Environ Model Software. 50:37–50. doi: 10.1016/j.envsoft.2013.08.012.
San-Miguel-Ayanz J, Moreno JM, Camia A. 2013. Analysis of large fires in European mediterranean landscapes: lessons learned and perspectives. For Ecol Manage. 294:11–22. doi: 10.1016/j.foreco.2012.10.050.
Scott JH, Thompson MP, Gilbertson-Day JW. 2017. Exploring how alternative mapping approaches influence fireshed assessment and human community exposure to wildfire. GeoJournal. 82(1):201–215. doi: 10.1007/s10708-015-9679-6.
Silva FRy, Molina Martínez JR, González- Cabán A. 2014. A methodology for determining operational priorities for prevention and suppression of wildland fires. Int J Wildland Fire. 23(4):544–554. doi: 10.1071/WF13063.
Syphard AD, Brennan TJ, Keeley JE. 2014. The role of defensible space for residential structure protection during wildfires. Int J Wildland Fire. 23(8):1165–1175. doi: 10.1071/WF13158.
Thompson MP, Freeborn P, Rieck JD, Calkin DE, Gilbertson-Day JW, Cochrane MA, Hand MS. 2016. Quantifying the influence of previously burned areas on suppression effectiveness and avoided exposure: a case study of the Las Conchas Fire. Int J Wildland Fire. 25(2):167–181. doi: 10.1071/WF14216.
van der Merwe M, Minas JP, Ozlen M, Hearne JW. 2015. A mixed integer programming approach for asset protection during escaped wildfires. Can J For Res. 45(4):444–451. doi: 10.1139/cjfr-2014-0239.
Viana-Soto A, Aguado I, Martínez S. 2017. Assessment of post-fire vegetation recovery using fire severity and geographical data in the mediterranean region (Spain). Environments - MDPI. 4(4):90. doi: 10.3390/environments4040090.

Weiss A. 2001. Topographic position and landforms analysis. In: 21st Annual ESRI International User Conference. San Diego, CA.

Wei Y, Thompson MP, Scott JH, O'Connor CD, Dunn CJ. 2019. Designing operationally relevant daily large fire containment strategies using risk assessment results. Forests. 10(4):311. doi: 10.3390/f10040311.

Wickham H. 2016. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag.

Wright MN, Ziegler A. 2017. ranger: a fast implementation of random forests for high dimensional data in C++ and R. J Stat Software. 77(1):1–17. doi: 10.18637/jss.v077.i01.

Zhou XH, Obuchowski NA, McClish DK. 2011. Statistical methods in diagnostic medicine [Internet]. John Wiley & Sons, Inc. doi: 10.1002/9780470906514.fmatter.