Leveraging national forestry data repositories to advocate wildfire modeling towards simulation-driven risk assessment

Juan Luis Gómez-González [a,*], Alexis Cantizano [a], Raquel Caro-Carretero [b], Mario Castro [a,c]

[a] *Institute for Research in Technology, ICAI School of Engineering, Universidad Pontificia Comillas, c/ Santa Cruz de Marcenado, 26, 28015 Madrid, Spain*
[b] *University Institute of Studies on Migration (IUEM), Chair in Disasters Fundación AON España, c/ Alberto Aguilera, 23, 28015 Madrid, Spain*
[c] *Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid 28015, Spain*

ARTICLE INFO

ABSTRACT

Modeling wildfire dynamics is complex and challenging due to the multiple scales involved in fire propagation, from physical–chemical processes to the interaction with topography and meteorological conditions. To provide reliable indicators of the risk of an ongoing wildfire, models aimed at informing policy-making should quantify the primary sources of uncertainty in their predictions. In this paper, we introduce a novel methodology built on top of Cellular Automata to assess the impact of uncertainty by implementing wildfire ensemble modeling using data from the Spanish National Forestry Data Repositories. Uncertainty is embedded in the model considering the $\pm 2\sigma$ deviations from the medians of linear regressions of the canopy stratum with LiDAR metrics as explainable variables. The relevance of dynamic meteorological conditions in contrast to static environment conditions is analyzed. Our results suggest that an accurate account of the fuel model, including time-dependent wind and moisture maps, is mandatory to provide reliable predictions. Using a real case study (Concentaina's extreme wildfire), we also illustrate the importance of assessing the impact of the firefighters' mitigation efforts.

1. Introduction

Wildfires in the Anthropocene era foster extreme events that surpass existing suppression resources (Bowman et al., 2020). Understanding their potential sources and researching their classification is necessary, as extreme wildfires are not isolated events (Oliveira et al., 2021). Nowadays, the likelihood of wildfire occurrence is estimated through indexes like the Fire Weather Index (FWI) (Canadian Forest Service Publications, 1978), which rates fire weather severity solely based on weather observations. Other indicators are commonly used in the context of wildfire risk assessment, such as the Keetch-Byram Drought Index (KBDI) (Keetch and Byram, 1968), the Fire Danger Rating System (FDRS) (Considine, 2006), or the Haines Index (Haines, 1988).

The complex dynamics of wildfire modeling are the outcome of many coupled physical–chemical processes (Liu et al., 2021; Sullivan, 2009a). They are addressed either by proposing *detailed models* that compute fire dynamics from fundamental principles (Wegrzynski and Lipecki, 2018; Pimont et al., 2009) or *operational models*, which rely on simplifications and empirical relationships between different fire behavior properties and the environmental conditions. Prominent examples of detailed models are the Wildland Urban Interface Fire Dynamics Simulator (WFDS) (Mell et al., 2007) and FIRETEC (Linn et al., 2002).

Alternatively, *operational models* aim to provide timely simulations shorter than real-time events, enabling effective management planning and rapid emergency responses (Sullivan, 2009b). They comprise two different modules: fire behavior and fire growth. The fire behavior module comprehends the simplification of the complex physical and chemical processes involved in wildfire combustion into a set of empirical or semi-empirical formulations. These rely on carefully calibrated environmental parameters based on laboratory-scale experiments. Besides, the fire growth module comprises a set of algorithms that propagates a given fire ignition over the spatial grid (Sullivan, 2009c). Propagation is based on the local rate of spread derived from the fire behavior properties. Over heterogeneous terrains and diverse

---

meteorological conditions, it leads to the simulation of complex fire perimeters.

However, their applicability is often limited to specific ranges of environmental conditions explored through experiments and observations. A comparative study of models involving 1238 observations (Cruz and Alexander, 2013) concluded that a 35% error in estimating fire spread rates is a valid standard for model acceptance. Another study testing the validity of models in prescribed burning experiments on single fuel beds reports a poorer performance of operational models compared to physical detailed models (Weise et al., 2016). Nevertheless, the study highlights that minor fuel parameterization adjustments can improve operative models' performance.

The benefits of operational models in terms of scalability, ease of use, and flexibility may overcome their limitations if they are correctly combined with high-quality/large-volume sources of information. In particular, remote sensing techniques incorporate sources of data such as Laser Imaging Detection and Ranging (LiDAR) (White et al., 2017), satellite observations (Seydi et al., 2022), availability of extensive weather information databases (Guo et al., 2023), and forestry inventories (González-Ferreiro et al., 2017).

However, the proper quantification of uncertainty in their predictions is a shared problem of both types of models. Detailed models require massive computation to fine-tune their parameters to capture any potential scenario. Operational models combine diverse sources of information with different degrees of reliability or granularity. A clear benefit of modeling uncertainty is the ability to manipulate which variables are more prone to provide sensitive variations in the final results (Benali et al., 2016).

In this regard, a potential solution is using the so-called ensemble modeling. This technique runs different simulations for a specific case study whose inputs are sampled from distributions. These distributions characterize the randomness present in different driver processes of the studied phenomena. In wildfire scenarios, the canopy structure of a forest (Kelly et al., 2017) or the meteorological conditions (Finney et al., 2011) are some examples. The set of the different simulation results represents a distribution of the possible outcomes. This approach's main advantage is the simplicity of generating confidence intervals from complex processes. In turn, these confidence intervals help to define new and more reliable indicators (Yousefi et al., 2020).

The importance of uncertainty quantification in advanced statistical analysis is a valuable tool in reducing the impact of climate change. In other fields, several challenging works have contributed to the search for mitigation measures to reduce the flooding risk (Lama et al., 2021, 2021b; Pirone et al., 2023), the assessment of soil sustainability by water erosion (Lense et al., 2023), the evaluation of land use and land cover to improve ecohydrology (Ray et al., 2023) or the use of different remotely sensed data to assess agricultural burned affected areas (Mohammad et al., 2023). Relevant to any ecological study, it is essential to highlight the value of both advanced experimental and modeling analysis of the prediction of natural phenomena (Lama and Crimaldi, 2021; Errico et al., 2019; Crimaldi and Lama, 2021).

In this work, we aim to bring together the advantages of ensemble modeling, operational methods, quantifying uncertainty, and utilizing reliable open data sources. Our goal is to develop a unique approach using Cellular Automata (CA), focusing on reliability and accuracy. This approach is built on top of pre-existing validated models such as FlamMap (Finney, 2006). A case study in Spain is used, the Concentaina wildfire, which is extensively documented and offers valuable scientific insights. This fire consumed a significant portion of the area within less than 24 h. Firefighting teams closely monitored the event, implementing suppression measures to contain its progression. Additionally, a simple modeling approach for incorporating suppression measurements is presented. The fire's characteristics, including its wind-driven nature and ignition, along with documented immediate fire suppression actions and detailed forest fire reports can help us understand post-fire dynamics. We chose this wildfire because the Spanish Mediterranean basin

is at high risk of wildfires, which is expected to increase (Arca et al., 2007). At the national level, Spain lacks the existence of an open public service that provides the necessary inputs for these models, in contrast to web services such as PREVINCAT for the region of Catalunya (González-Olabarria and Piqué, 2019) or LANDFIRE in the United States (Reeves et al., 2009). We compensate for this by proposing to the community a data processing framework to produce this information using sources such as a nationally coordinated LiDAR repository and forest inventory surveys.

In Section 2, we provide a detailed explanation of our methodology, the sources of information used, the modeling framework, and its physical parameterization. Section 3 details the fire event of Cocentaina 2012. Our findings are described in Section 4, and we conclude with a discussion of the potential benefits of our approach in Section 5.

2. Materials and methods

The methodology developed in this work aims to be flexible and straightforward to deploy, leveraging the availability of public data in the Spanish context. This section explains how the essential input magnitudes are derived from public information repositories. It also details the modeling techniques and processes used to capture uncertainties. Fig. 1 displays the elements that integrate the workflow of this research.

2.1. Public data repositories

Three major public forestry repositories are considered in this study. First, the LiDAR survey conducted under the *Spanish National Orthophoto Program* (PNOA) (Spanish National Geographic Institute IGN, 2008) provides valuable information on the forested areas. Second, the *Spanish Forest Map* (MFE) (MITECO, 2005) offers a comprehensive tessellation of soil use and categorization across Spain. Lastly, the public *Spanish National Forest Inventory* (IFN) (MITECO, 1997) contributes with detailed data on forest properties. Detailed information can be found in Appendix A. Geographical information present in the information from these databases is processed and visualized in QGIS (QGIS Development Team, 2021), a Geographical Information System (GIS) software.

2.2. Cellular Automata: fire behavior and wildfire growth model

Two operational models widely accepted are Prometheus (Tymstra et al., 2010) and FlamMap (Finney, 2006), created by the Canadian Forest Service and the US Forest Service, respectively.

In particular, FlamMap Version 6.1 incorporates two different fire growth modules: Minimum Travel Time (MTT) (Finney, 2002) and Fire Area Simulator (FARSITE), based on Huygens' Principle for fire front spread (Finney, 1998). MTT assumes background fire behavior conditions to be static. FARSITE accommodates time-varying environmental conditions and incorporates some dynamic features of fire behavior.

Both approaches use the same fire behavior model, which follows Rothermel's model (Rothermel, 1991). This theoretical framework dates back to the 1970s and is extensively used, constituting the core of most *operational models*. Subsequent improvements have been further introduced, such as modules that incorporate crown fire transition (Wagner, 1977; Rothermel, 1991), crown fire rate of spread (Finney, 1998; Scott et al., 2005), fuel moisture modeling (Albini, 1976; Nelson, 2000), among others.

However, the MTT and Huygens' wavelet principle for fire perimeter expansion have some drawbacks. MTT does not allow variable environmental conditions, such as whenever the fire lasts longer than the time scale over which meteorological changes occur. On the other hand, Huygens' principle allows dynamical conditions but presents mathematical artifacts when fire fronts overlap. This procedure introduces an extra source of computation time and fire growth uncertainties. Under constant conditions, MTT and FARSITE yield equivalent results (Finney,
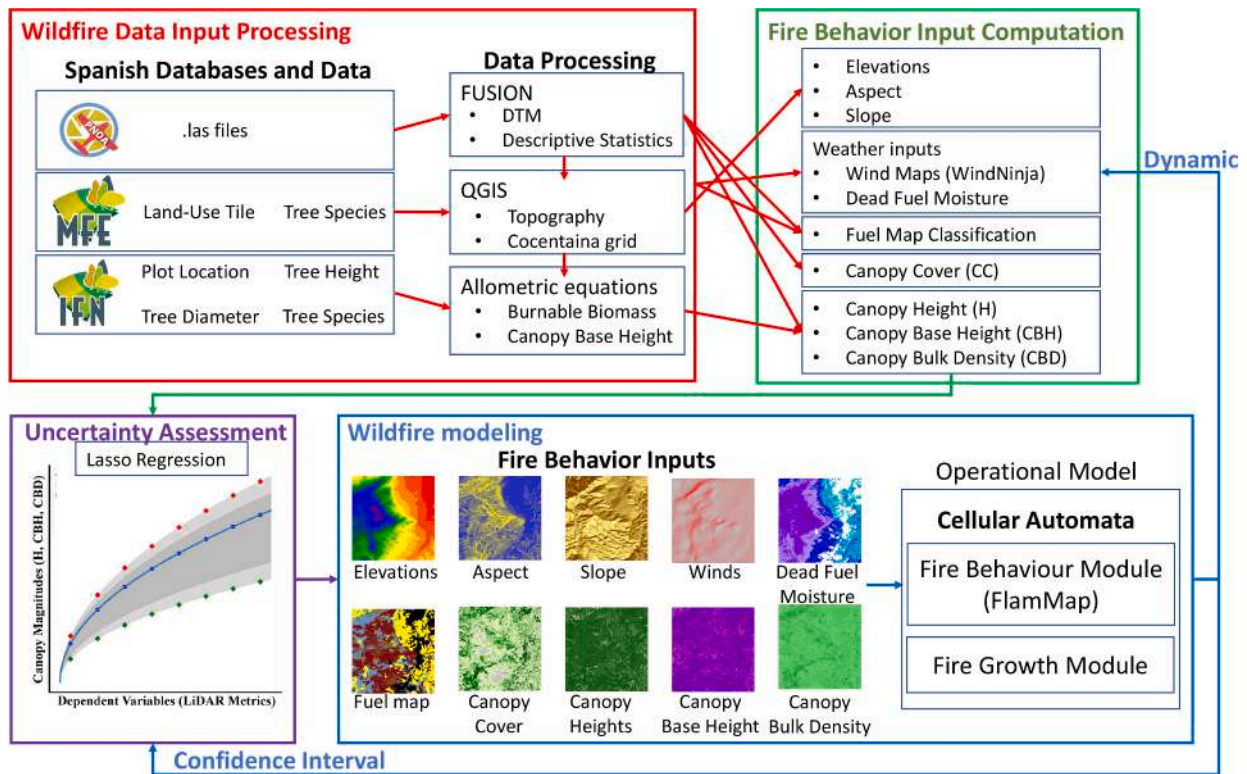
Fig. 1. Research flowchart. This work introduces a comprehensive methodology for modeling wildfire events in Spain using data from public repositories. The methodology results in an operational wildfire model integrating fire behavior and fire growth modules. Ensemble modeling techniques are applied to enhance predictive accuracy, allowing the presentation of results as confidence intervals. This approach involves modeling uncertainties in canopy variables sourced from forestry data. This framework considers dynamic environmental conditions and incorporates basic suppression modeling, ultimately leading to improved predictions of burned perimeters.

1998), illustrating that fire growth algorithms can be diverse yet reproduce the same outcome. Both methodologies run faster than real-time events, taking minutes to simulate hours of fire.

Another flexible propagating scheme is Cellular Automata, which, in the context of wildfire modeling, can be employed to simulate the spread of fire across a landscape (Purnomo et al., 2021; Mutthulakshmi et al., 2020; Trucchia et al., 2020). One of the advantages of this methodology is that it easily allows coupling fire behavior models with other non-physical dynamics, such as fire suppression modelingAlexandridis et al., 2011).

We introduce a CA model to maintain the robustness of MTT models and improve their flexibility. CA are dynamic frameworks that imply several advantages, such as:

- They naturally accept time-varying landscapes by pre-computing the expected conditions at the times required.
- Fire growth can be equivalent to MTT or Huygens' principle by adequately embedding fire behavior into the interacting grid.
- Fire front interactions are naturally tackled, with no extra requirements for mathematical artifacts.
- The CA models fuel combustion by a series of discrete states. For instance, un-burned, burning and burned. Other states can be envisaged.
- Computational costs increase with system size, the number of close neighbors, and the number of states of the grid elements, independently of the fire behavior modules considered. This makes it easier to make further improvements or changes.

The proposed CA uses the fire behavior module in FlamMap to estimate the fire spread rates. This semi-empirical model correlates the spread rate of a wildfire to critical environmental variables. These variables include wind speed, terrain slope, fuel particle size, moisture, and heat content. It is further related to the simple shape of a fire perimeter from point-source combustion. This shape is commonly approximated to be elliptical, following community consensus in the field.

Our CA is a deterministic model, meaning its evolution from a specific set of initial states is always similarly prescribed. The landscape is partitioned into a regular, squared lattice with cells of size $s \times s$. These grid elements are characterized by un-burned and burning states, representing 0 and 1, respectively. The spread of fire from cell to cell is facilitated by the close neighbor structure prescribed by a Moore neighborhood. Under this schemet each cell (i, j) communicates with its eight nearest neighbors on the grid. At the domain limits, close boundary conditions are applied. To translate the fire spread rates, the first step is to consider that fire propagates in all directions, following an elliptical distribution of velocities over spatial directions. Therefore, spread rates projected on the directions connecting center-center cells from these ellipses are considered, given the lattice's regular structure. This results in an interval of time Δt for fire to spread between cell centers, which is given by $\Delta t(i,j) = s_{ij}/k(\theta_{ij})$. A $\sqrt{2}$ factor is considered for diagonal distances due to the square geometry. This approach is illustrated in Fig. 2.

In this study, the CA methodology involves pre-computing the time a fire front takes to spread from any given cell to its closest neighbors under static environmental conditions. For a grid lattice comprising $n \times n$ elements, the total number of interacting "fire spread paths" is $8 \cdot (n-2) \cdot (n-2)$, $5 \cdot 4 \cdot (n-2)$, and $4 \cdot 3$ for inner cells, border cells, and corner cells, respectively. The time step Δt to integrate time is fixed in the model. The information that is tabulated and stored in memory when the simulation starts is the number of fixed time steps n that a fire front takes to spread from a cell i to another j, i.e., $n(i,j) = \Delta t(i,j)/\Delta t$. In this

$$\Delta t_{i-1,j+1} = \frac{S_{i-1,j+1}}{k_{i-1,j+1}} \quad \Delta t_{i,j+1} = \frac{S_{i,j+1}}{k_{i,j+1}} \quad \Delta t_{i+1,j+1} = \frac{S_{i+1,j+1}}{k_{i+1,j+1}}$$

$$\Delta t_{i-1,j} = \frac{S_{i-1,j}}{k_{i-1,j}} \qquad \Delta t_{i+1,j} = \frac{S_{i+1,j}}{k_{i+1,j}}$$

$$\Delta t_{i-1,j-1} = \frac{S_{i-1,j-1}}{k_{i-1,j-1}} \quad \Delta t_{i,j-1} = \frac{S_{i,j-1}}{k_{i,j-1}} \quad \Delta t_{i+1,j-1} = \frac{S_{i+1,j-1}}{k_{i+1,j-1}}$$
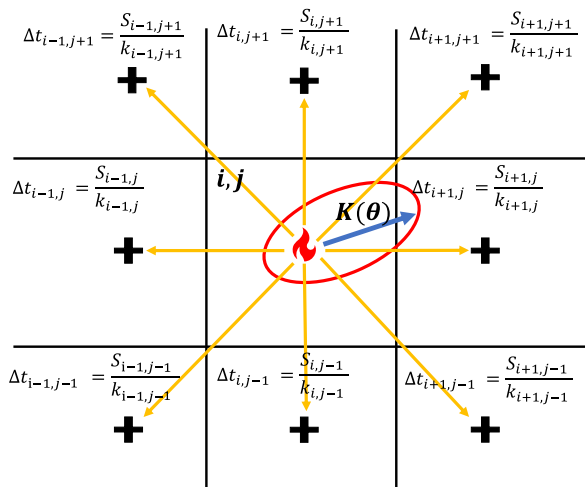
Fig. 2. The time intervals that the deterministic CA takes to spread from a cell to its neighbors. Spread rates correspond to projections of the elliptical shape profile in the directions of the connecting neighbors.

model, the non-burning cells constituting the boundary of the fire perimeter are the elements susceptible to ignite after a CA iteration. The CA integrates an amount of time $t = t + n_{min} \cdot \Delta t$ per iteration. The amount of time $n_{min} \cdot \Delta t$ is equivalent to the smallest number of fixed time steps the fire front takes to spread and ignite a subset of the non-burning boundary cells. The quantity n_{min} corresponds to the minimum of $n(i,j)$ at the boundary of the fire perimeter. n_{min} can be shared among several fire-spreading paths, thus simultaneously igniting several cells at the burning front perimeter in a single CA iteration. Continuous fire front expansion is considered in the case it has not reached neighboring centers by initializing particular counters, $n^*(i,j) = n(i,j)$. After each CA iteration, the number of time steps required for fire to spread is updated to $n^*(j,i) = n^*(j,i) - n_{min} \cdot \Delta t$. Together with the new susceptible cells to be burned, these values are examined for the next iteration and continuously updated until complete fire propagation. The value of Δt must be smaller than any of the intervals of time fire takes to traverse existing fire spread paths, i.e., $\Delta t \leqslant \min \Delta(j,i)$. This is necessary to avoid physically inaccurate over-spreading. Fig. 3 illustrates this mechanism.

2.3. Model parameterization

Rothermel's model estimates fire spread rates employing a set of variables easily interpreted from the environment. Scientists and stakeholders can retrieve this data over large areas using remote sensing
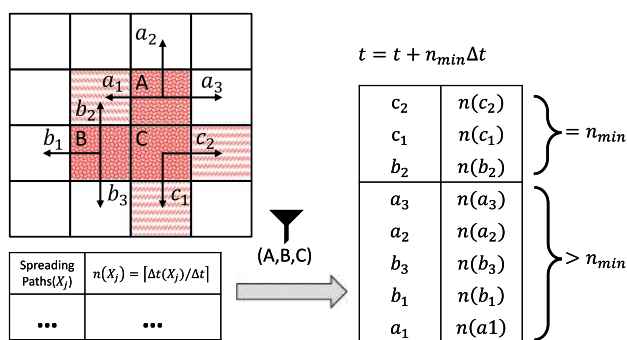

Fig. 3. CA iteration scheme. The precomputed Δt (a fire front spreads from a cell to its closest neighbors) is filtered only to include ignited cells fire spreading paths. The example shows ignited cells A, B, C and their fire spreading paths a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2. The same minimum number of steps is shared by b_2, c_1, and c_2. The system iterates n_{min} time steps and ignites corresponding cells. For visualization purposes, diagonal interactions are omitted.

techniques and process it into appropriate databases. The study area's spatial characterization relies on raster layers, including:

- Topographic information (elevation, slope, and aspect).
- A fuel model map (combustible properties).
- Canopy structure variables (canopy cover, stand height, canopy base height, and canopy bulk density).
- Wind data (speed and direction).

Moreover, accurate meteorological information is crucial for pre-conditioning the dead fuels' humidity.

This work completes the derivation of these magnitudes using three public data repositories specific to the Spanish context. While tailored to Cocentaina's wildfire scenario, the approach can be adapted to other Spanish regions with coherent wildfire data. It is worth noting that the accuracy of forestry mapping may be compromised after a wildfire occurs in the same area. Likewise, data from different forestry databases may not match due to being updated unevenly after extensive periods of time. These inherent limitations and considerations are fundamental for wildfire modeling, requiring continuous evaluation and refinement.

2.3.1. Topographic information

Three topographic magnitudes are required to characterize the model: elevation, aspect, and slope. The elevation map corresponds to the Digital Terrain Model (DTM) of the wildfire scenario area, readily available from LiDAR information. In this study, point cloud data is derived from the first coverage of the PNOA (recorded in the year 2009), using data batches 3 (covering Castilla la Mancha, Murcia, Alicante) and 8 (covering Valencia). These LiDAR datasets are publicly available from the *Spanish National Geographical Institute* (IGN). The LiDAR data is processed using LAStools (rapidlasso GmbH, 2021) and further analyzed using FUSION software (McGaughey, 2022).

To this extent, the LiDAR cloud points that overlap the modeled wildfire area are processed into a 25×25, m^2 DTM of 321x321 elements (covering an area of 8×8 km^2). This DTM is converted into raster information ready to be implemented in the fire behavior module. Aspect and slope layers are estimated from the elevation raster map by the appropriate tools in QGIS. Fig. 4 shows LiDAR cloud points considered in this study within the province of Alicante, as displayed on a satellital map.

2.3.2. Fuel models

In the specific case of Alicante, the *Department of Agriculture, Rural Development, Climatic Emergency, and Ecological Transition* developed a fuel mapping based on Scott et al.'s fuel classification in 2019 (Scott et al., 2005). However, this information is available only after the wildfire event. Therefore, a fuel map is developed here by combining the Spanish Forest Map at scale 1:50000 (MFE50) land use information and LiDAR data descriptive statistical metrics, following the guidelines outlined in (Sánchez García et al., 2019). A more updated version of the Spanish Forest Map at scale 1:250000 (MFE25) is currently in development. However, the progress of MFE25 implementation varies across different provinces. While MFE50 is accessible for the entire Spanish geography, MFE25, for instance, is not yet available in Alicante.

Concerning the MFE50 data, the vector file corresponding to the Community of Valencia, which includes the province of Alicante, undergoes processing in QGIS. Thus, the MFE50 polygonal information is intersected with a geometrical grid (25×25, m^2 grid: 321×321, cells) representing the Cocentaina area modeled in this study. This intersection retains the relevant information from the MFE tiles.

A decision tree is employed to associate the fuel models (Fig. 5), which relies on two aspects. Firstly, on specific details from the MFE, cells are considered as: Non-Combustible (NC), dense Forests (F), Shrublands (S), and Grasslands (G). Secondly, on a set of structural descriptive metrics of the fuel stratum: total Vegetation Height (VH), Shrubland Height (SH) and Shrubland Cover (SC).
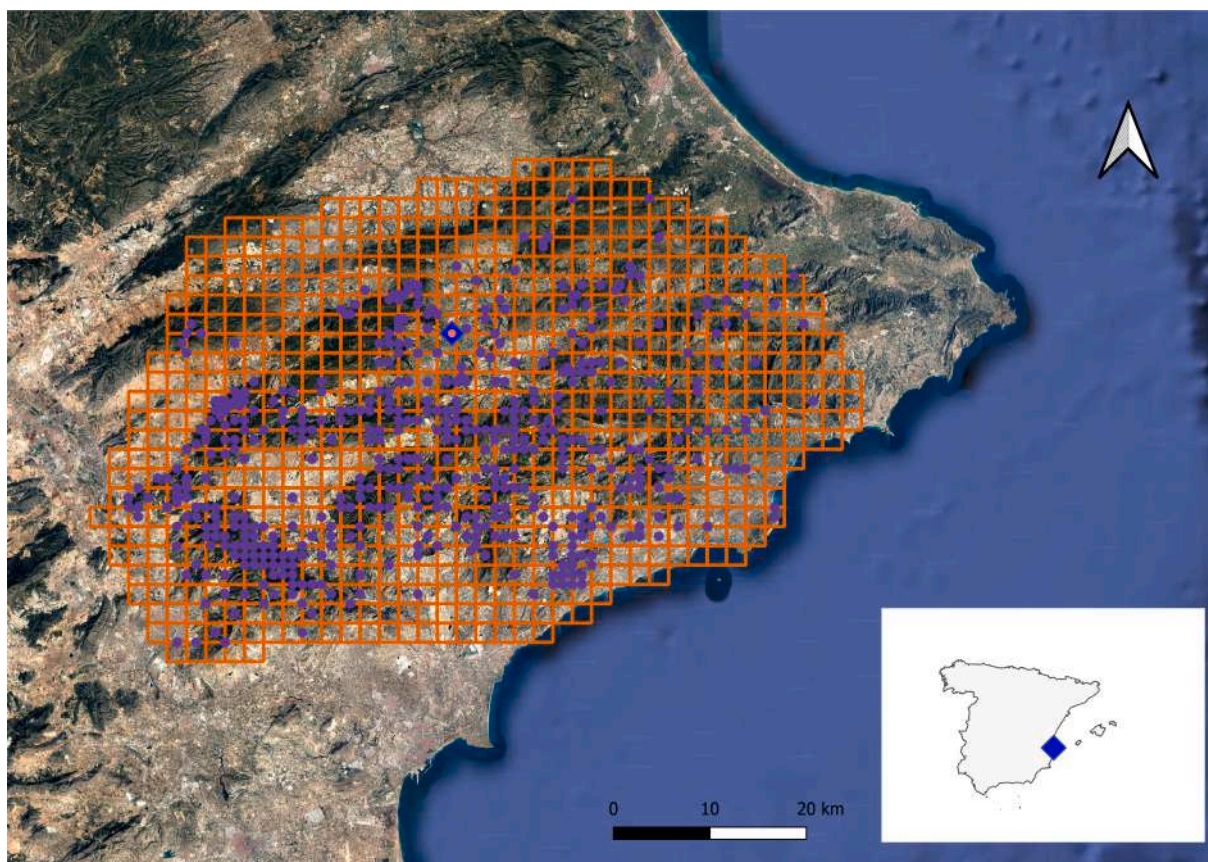
Fig. 4. Geographical surface of the area surrounding Cocentaina. The diamond marker on the map indicates the location of Cocentaina municipality. The grid polygon illustrates LiDAR data tiles from the PNOA employed as input variables to train the three canopy structure models. Circular markers correspond to the location of the IFN3 plots used as target variables in the canopy structure models. Inset map was created using geoBoundaries (Runfola et al., 2020).
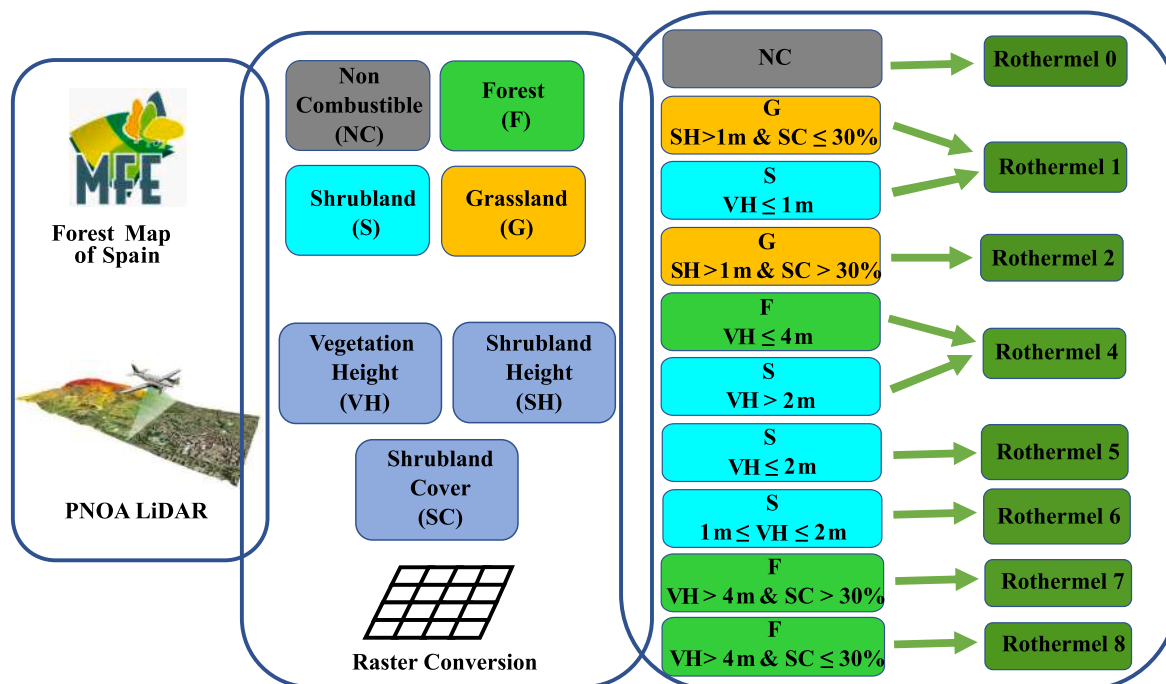

Fig. 5. Decision tree that assigns different Rothermel fuel models considering PNOA LiDAR metrics and soil use information from the MFE.

The vegetation structural descriptive metrics are associated with statistical metrics of the LiDAR cloud points normalized to the ground level defined by a DTM. This computation employs a DTM with a resolution of 2×2 m^2 to normalize LiDAR data within the 25×25, m^2 grid elements of the simulated region. Therefore, VH is associated with the 90th percentile of the return cloud points from a LiDAR dataset that restricts the heights of these reflections from 0.5 to 40 m (40 m spans the whole vegetation height distribution). On the other hand, the shrub metrics consider a LiDAR dataset that restricts the heights of these reflections from 0.5 to 2 m (a height range that approximately only includes the shrub stratum). Then, SH is associated with the mean height of the return cloud points, and SC is related to the percentage of first returns above 0.5 m. These statistical metrics are computed using FUSION.

2.3.3. Canopy structure

The canopy structure encompasses the vertical fuel distribution above the surface fuel models, which are characterized in Section 2.3.2. Thus, canopy structure refers mainly to the biomass contribution of trees in the forest cover. Deriving relevant magnitudes is complex as the forest cover is heterogeneous in tree species, morphology, and spatial distribution. Moreover, conducting local measurements requires large investments to comprehend large domains. As a more affordable approach, local in situ observations constitute the information to train regression models that extrapolate the canopy structure using data collected by remote sensing techniques.

The methodology to characterize the canopy structure is based on (Wagner, 1977). Each grid element is assumed to have a homogeneous canopy structure. Four input layers to FlamMap's fire behavior module are designed to represent this information and are ubiquitous in *operational models*: Canopy Height (H), Canopy Base Height (CBH), Canopy Bulk Density (CBD), and Canopy Cover (CC). The average tree height estimates H, and the average tree surface-to-crown height estimates CBH. The resulting value from the average superficial density of aerial fuel biomass divided by the average crown height estimates CBD. The procedure to retrieve the CC layer is simpler as LiDAR data is a direct estimator of this magnitude (Arumäe and Lang, 2018), in particular the percentage of first returns above 2 m up to 40 m.

There exist approaches that consider individual tree morphology to compute H, CBH, and CBD as in (Fidalgo-González et al., 2019). However, they are less flexible due to the broad range of tree species and the lack of adequate allometric models to estimate tree biomass/geometry. Other inaccuracies arise because the PNOA LiDAR and IFN coverages do not overlap in time. Consequently, their information may not match or fully represent a wildfire as the time between data collection and the fire event is of the order of years. Similarly to the case of the most updated revision of the MFE, the 4th Spanish National Forest Inventory (IFN4) data for Alicante is still being updated and unavailable. Therefore, the 3rd Spanish National Forest Inventory (IFN3) database is utilized for this study. The relevant information from this inventory comprehends the geographical position, tree species, tree heights, and trunk diameters. Major trees are referenced radially from the center of the plot up to a radius of 25 meters.

Implementing the canopy structure involves the creation of regression models using information from the IFN3 and LiDAR data from the first coverage of the PNOA as input values.

Initially, only those IFN3 plots contained in the study area of the PNOA LiDAR coverage data are considered, as can be seen in Fig. 4. It resulted in a total number of 876 plots. Then, the MFE50 is used to extract the predominant tree species in the 8×8, km^2 grid that constitutes the simulated area of Cocentaina. *Pinus Halepensis* stands out as the predominant species. There is a minority presence of 4 other species: *Pinus Pinaster, Populus (nigra x Canadensis), Juniperus Oxycedrus*, and *Quercus Ilex*. Thus, restricting the initial set of IFN3 plots to those with these species, only 492 plots remain, as shown in Fig. 4.

Next, the H, CBH, and CBD values that characterize the IFN3 plots are calculated. H is obtained directly by averaging the value of heights. Estimating CBH requires knowledge of allometric equations for the crown height or the surface-to-crown height of the trees. However, no appropriate models of the predominant tree in the Cocentaina Area (*Pinus Halepensis*) were found in the literature. In the absence of other models better suited to the fire context, the species *Pinus Pinaster*, which is phylogenetically close to *Pinus Halepensis*, is extensively characterized in (González-Ferreiro et al., 2017). Thus, its allometric expression for CBH from (González-Ferreiro et al., 2017) is implemented. The combined presence of *Pinus Halepensis* and *Pinus Pinaster* completely dominates the canopy. Thus, the contribution from the rest of the trees is neglected to compute the CBH. For the CBD layer, only the contribution of aerial biomass susceptible to burn in crown fires is considered. It includes particles of 2 cm diameter size or smaller, leaves, and needles, which agrees with the characteristic distribution of aerial fuel particles burning at the fire front (Stocks et al., 2004; Sando et al., 1972). Aboveground fuel biomass allometric models for species in Cocentaina as a function of tree diameter and height are reviewed in (Montero and Ruiz-Peinado, 2006). Dividing the sum of total biomass per plot by the area of the plot is an estimate of the crown fuel load (CFL) or surface density of aerial fuels. Finally, CBD consists of dividing CFL by the crown size estimates of the plot (the difference H-CBH).

Finally, the different LiDAR statistics are computed geographically restricted to the IFN3 plot areas. They are expected to be adequate estimators of forest canopy since they are referred to the height distribution of the LiDAR point clouds. The IFN3 LiDAR dataset presents high dimensionality since up to 41 height statistical metrics (this study ignores laser intensity metrics) returned by FUSION are considered descriptive variables of the canopy information. Moreover, many metrics are redundant, so high collinearity is expected. On the other hand, the canopy magnitudes obtained from the IFN3 plots present a high noise level, given the uncertainties present in the data and the various allometric models used.

2.3.4. Canopy uncertainty assessment

A dispersion analysis in the canopy data shows it does not follow a Gaussian distribution, as shown in Fig. 6; so, heteroscedasticity in the fitting residuals is expected. In order to properly model the stochastic functional relationship between IFN3 data and estimated quantities, we need to address the heteroscedasticity problem and then calibrate the mathematical model. Therefore, the proposed modeling procedure is based on two decisions: First, a Box-Cox transformation on the target variables and a Yeo-Johnson transformation on the input variables are applied so that the distribution of the residuals is homoscedastic (Sakia, 1992). Next, a linear Lasso regression is applied to mitigate the problem of collinearity in the input data, as it performs variable selection during fitting (Zou and Hastie, 2005). The regularization parameter is optimized using a grid search cross-validation scheme. Once the fitting is completed, observations beyond 3σ values are omitted to eliminate the influence of outliers. Then, the fitting process is repeated without them. The transformation of the variables and the linear fitting are implemented using the scikit-learn Python library (Pedregosa et al., 2011).

Three linear regressions are performed after the Yeo-Johnson transformation for H, CBH, and CBD and the Lasso regression technique identifies the following relevant features:

- H as the predictor: minimum (light return) height, height variance, coefficient of variation, height kurtosis, 10th and 90th height return percentiles, and profile area.
- CBD as the predictor: height variance, coefficient of variation, interquartile distance, skewness, kurtosis, median of the absolute deviations from the overall median, 1st percentile value, 5th, 30th, 70th, 90th, 95th, and 99th height return percentiles, percentage of first returns above the mean, percentage of all returns above the norm, and profile area.
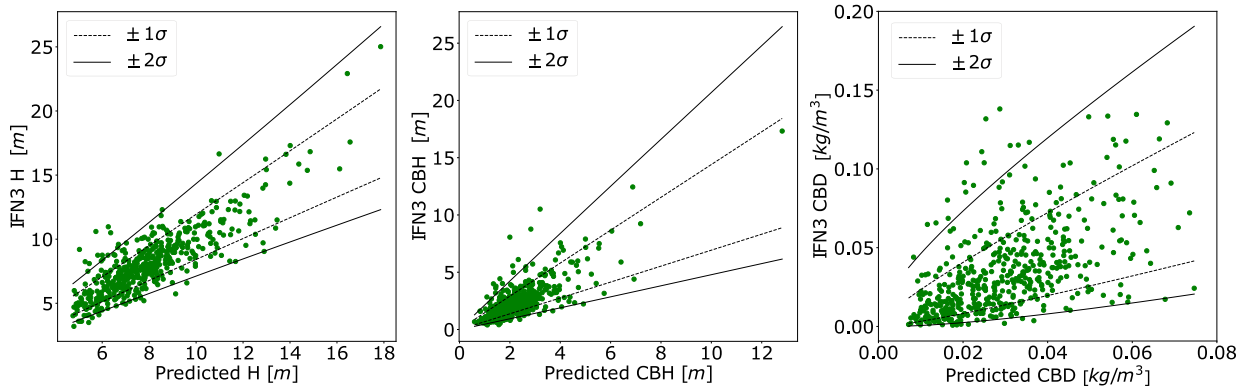
Fig. 6. H, CBH, and CBD magnitudes from the 492 IFN3 plots considered in this study versus the corresponding regression results. Regression estimates of the standard deviations at $\pm\sigma$ and $\pm2\sigma$ are plotted against data points. As can be seen, the deviations are not constant for the different ranges of values; therefore, it is a heteroscedastic model of the noise in the data.

- CBD as the predictor: minimum (light return) height, mode of the return distribution, variance, coefficient of variation, kurtosis, 5th, 75th, and 80th height return percentiles, canopy relief ratio, percentage of first returns above 2 m, and percentage of first returns above the mean.

Extrapolating these regressions to the study domain requires applying the inverse transformation. Thus, a heteroscedastic uncertainty model of the forest cover layers is obtained based on the uncertainty associated with the available raw data. Besides, the extrapolation requires deriving the statistical metrics of the LiDAR return cloud points corresponding to the 8×8, km^2 grid surface elements that model the Cocentaina area. These specifications are required since the input LiDAR metrics need to be derived from a normalized point cloud with the exact

resolution used for the input data processing. Regression results from the IFN3 data, along with uncertainty levels, are presented in Fig. 6. As shown, the method of transforming (Yeo-Johnson) regression and inverse transforming captures the variability of the data across all values.

With these calibrated models, we can now generate scenarios with different degrees of uncertainty and propagate that uncertainty into the fire model. Fig. 7 illustrates the methodology for deriving perturbed forest cover layers according to uncertainty levels compatible with the input data.

2.3.5. Wind maps

Fire behavior is significantly influenced by wind speed and direction. Fire spread predictions improve as more realistic wind fields are computed (Forthofer, 2007). Computational Fluid Models (CFD) can offer accurate approximations of the wind fields, as direct measurements are not readily available at the necessary level of resolution. Nonetheless, their computational demands often exceed the requirements of fast-fire behavior modeling approaches.

FlamMap incorporates the operative wind solver WindNinja (Wagenbrenner et al., 2016). The solver consists of a simplified tridimensional CFD model conceived to compute wind field maps involving simulation times of the order of minutes or less. Given a specific wind speed and direction observation, it reduces the modeling approach to be consistent with mass-conservation and topography. This integration enables a compatible and practical solution for operational wildfire spread modeling. It has to be mentioned that meteorological observations in wildfire contexts are typically limited to specific stations or local monitoring.

2.3.6. Live and dead fuel moisture

Humidity is a relevant variable in fire behavior, with ongoing developments in remote sensing techniques that aim to estimate its spatial distribution (Myoung et al., 2018). Here, humidity is modeled based on moisture dynamics. FlamMap integrates the estimation of dead fuel moisture values following the approach proposed by (Albini, 1976 and Nelson, 2000) (Fig. 9 right). The approximation uses preceding meteorological conditions (temperature, relative humidity, precipitation, and cloud coverage), topography, and tree canopy to precondition moisture in equilibrium with the environment. Moisture preconditioning takes place for each fire behavior map that is computed. The stream of meteorological data used for each update is the full-time series up to the instant of time of interest. Initial values for moisture ratios are required to solve this equilibrium. For reproducibility purposes, initial valid estimates are indicated. For fuel models, these are assigned identically. Moreover, these depend on the time lag that different combustible particle sizes take to respond to environmental changes. 6% for 1 h, 8% for 10 h, 10% for 100 h (Fig. 8). The herbaceous fuel (125%), and live
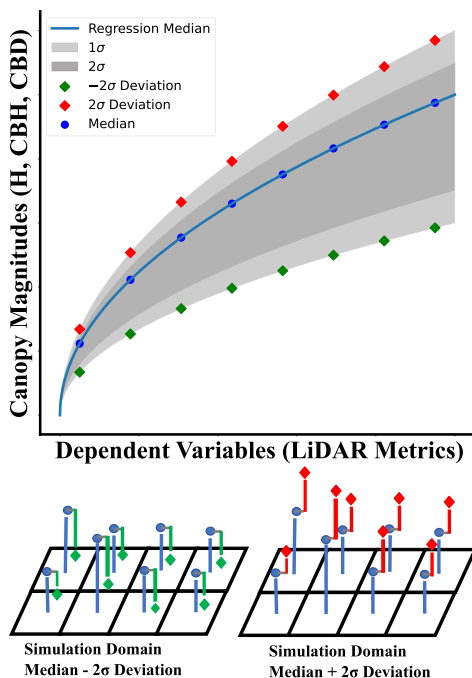

Fig. 7. Schematic diagram of the methodology to use LiDAR data-driven models of H, CBH, and CBD with uncertainty retrieval. Lasso regression is used to fit the model. This methodology enables a robust estimation of uncertainty and subsequent perturbation of the canopy input layers. Sampling values from the uncertainty distribution enables the implementation of ensemble modeling techniques.
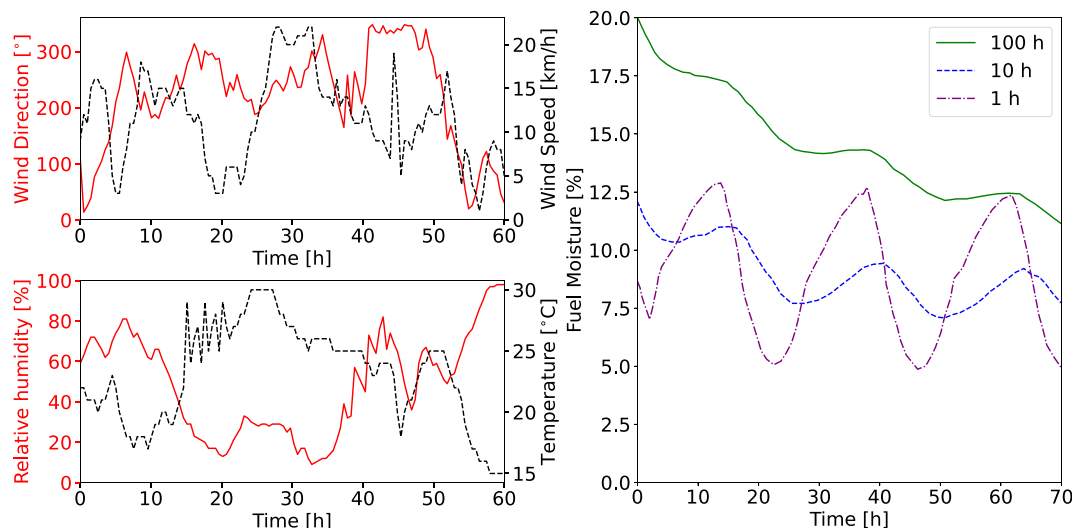
Fig. 9. Left, meteorological data during the wildfire event. Records presented in this figure start at 1:10 pm on 12th July 2012, over 60 h. Relative humidity, temperature, wind speed, and wind direction are displayed (Soriano and Botella, 2015). Right, an example of dead fuel moisture computation over 70 h. This example illustrates a fuel composed of wood sticks. Moisture dynamics is presented for three different particle diameters. The legend denotes the diameter by their time lag in hours (h) to respond to environmental conditions. The smaller the diameter, the faster the fuel achieves stationary equilibrium with environmental conditions. Computation reproduced from (Nelson, 2000).

Fig. 8. Aerial photos of the burned surfaces. Source (Soriano and Botella, 2015).

woody fuels (100%) are also considered. In addition to surface fire spread modeling, crown fire modeling requires the specification of foliar moisture content. Foliar moisture content is set to 100%. This value is not subjected to moisture preconditioning.

3. Case study: Cocentaina wildfire

The case study refers to a well-documented wildfire on July 12th, 2012, in the outskirts of Cocentaina, a small town in the province of Alicante, southeast Spain.

The forested area corresponds to the natural protected area known as "Parque Natural de la Sierra de Mariola." The dominant canopy species found are Aleppo Pine (*Pinus Halepensis*) and Oak (*Quercus Ilex*). The study area exhibits a typical Mediterranean landscape, with some resemblance to a continental climate due to the high altitude of the mountain range. Winters are cold, with temperatures dropping as low as −15 °C at the highest peaks (around 1000 m of altitude). Summers are hot, with temperatures rising above 35 °C and occasionally reaching 40 °C. Rainfall ranges from approximately 350 mm to 900 mm annually and follows highly irregular patterns.

The information presented here is compiled from post-fire reports available at the *Integrated Forest Fire Management System of Generalitat Valenciana* (SIGIF) (Soriano and Botella, 2015). As a significant contribution to this work, a comprehensive effort has been made to incorporate suppression activities. They are compiled from the corresponding

Forest Fire Report (PIF) according to the *General Statistics on Forest Fires* (EGIF) (MITECO, 2012), which is filled out by the firefighting authorities after the wildfire event. Further modeling of the terrestrial measurements is implemented following detailed reports in press (Shakila et al., 2012).

The wildfire started at 3:10 pm on July 12th and was entirely controlled by 8:15 am on July 16th. By 2:30 pm on July 13, the wildfire had already affected 90% of the final total burned surface. Wind primarily drove this wildfire and benefited from favorable upward topographical alignments. Initially, the fire propagated upwards following the canyon called "Barranco del Bou". Then, as the wind direction shifted westwards, it contributed further to the fire spread. The ignition source of the fire was located in a chaparral area with a high fuel load. The bulk of the fire perimeter exhibited a heterogeneous distribution of fuels, including areas with adult trees experiencing crown fire activity. Surface firefighters arrived 40 min after the fire was spotted. During the following 24 h post-ignition, up to 20 aircraft and 200 personnel worked to reduce fire spread on different fronts.

Meteorological conditions were recorded at the Agres Meteorological Station (CEAMET), located 8 km away from the fire site. At the start of the fire, temperatures were mild during the daytime (around 23 °C) and did not drop below 20 °C at night. The following day, temperatures increased to around 30 °C. Relative humidity varied between 60% and 80% during the first day but abruptly dropped to values ranging from 20% to 30% during the night (Fig. 9 left). The final fire perimeter

covered an area of 20.83 km with a total burned surface of 545.93 hectares in the forestry area. There were no secondary ignitions, and no urban or industrial interfaces were affected.

4. Results

Half-hourly spaced information on temperature, relative humidity, wind speed, and wind direction recorded from July 9th to 14th at the nearby Agres Meteorological Station constitutes the meteorological time series data. Within the dynamic wind model configuration, the input wind speed and direction in WindNinja are derived from the average of wind observations recorded over two hours. In every simulation, moisture is preconditioned using data from July 9th to the actual time fire behavior is being computed. Fuel Models, canopy cover, elevation, slope, and aspect maps are static throughout the different simulations. The impact of uncertainty in the canopy structure is estimated using the confidence interval $\pm 2\sigma$ concerning the median of the regression models of the three canopy layers H, CBH, and CBD. According to the model, this range encompasses 95% of the expected canopy layer deviations.

The final burned area is measured in hectares. The over and under-predicted burned areas, expressed as percentages relative to the actual burned surface, are presented as wildfire model mismatch estimators. It is crucial that as time passes and weather changes, the CA must implement these dynamics in an operational model, even though the effects of fire on local weather are neglected. Varying environmental conditions are considered by computing fire behavior maps at different times. The CA spreads the wildfire under constant conditions for each new fire behavior map for a certain number of time steps Δt. The temporal evolution of the landscape is simulated by utilizing the final state of the CA simulation as the initial condition for the subsequent set of environmental conditions.

FARSITE, integrated within FlamMap, adopts the same approach for simulating fire spread. However, it introduces additional complexity compared to the CA dynamics by representing the fire perimeter as a collection of geometric points. These mechanisms require continuous self-evaluation of the growing burned rim to prevent spurious overlapping. In contrast, the CA dynamics simulate burned surface growth without the need for self-evaluating mechanisms, as interactions between cells are defined and pre-computed before the simulation starts.

Fire suppression actuation is not part of the fire behavior module. Its appropriate modeling would require a detailed spatial and time distribution of their deployment. Press reports (Shakila et al., 2012) described where primary fire attacks occurred during day and night. Fire surface suppression is incorporated in the model by delimiting barrier elements on the grid where fire spread is not allowed. The extent of the coarse-inferred barriers is displayed from Fig. 10 to Fig. 13. Simulations presented in these figures take the median H, CBH, and CBD values from the canopy model regressions as inputs to the fire behavior module. Numbers highlighted next to the barrier layouts represent the time sequence of their execution:

- Barrier 1 was deployed at 15:45 (12th).
- Barrier 2 was deployed at 19:30 (12th).
- Barrier 3 was deployed at 21:30 (12th).
- Barrier 4 was deployed at 02:15 (13th).

According to the PIF and post-wildfire report, it took several days to completely control the fire, even though the burned perimeter remained relatively unchanged after 2:30 (13th). The fire spread mechanism does not incorporate natural fire extinguishing mechanisms. This introduces an extra layer of uncertainty in the modeling process. Thus, two finishing scenarios are explored: at 2:30 and 6:30, representing 14 and 17 h, respectively. Fire behavior is updated approximately at intervals of $2 \sim 3$ h, and $\Delta t = 40\ s$ is used, which is optimized for this scenario according to subSection 2.2.

In our case, we include four different fire behavior modifications:

- The static case equivalent to FlamMap with moisture preconditioning and WindNinja winds.
- The scenario in which wind changes dynamically.
- The scenario in which moisture preconditioning changes dynamically.
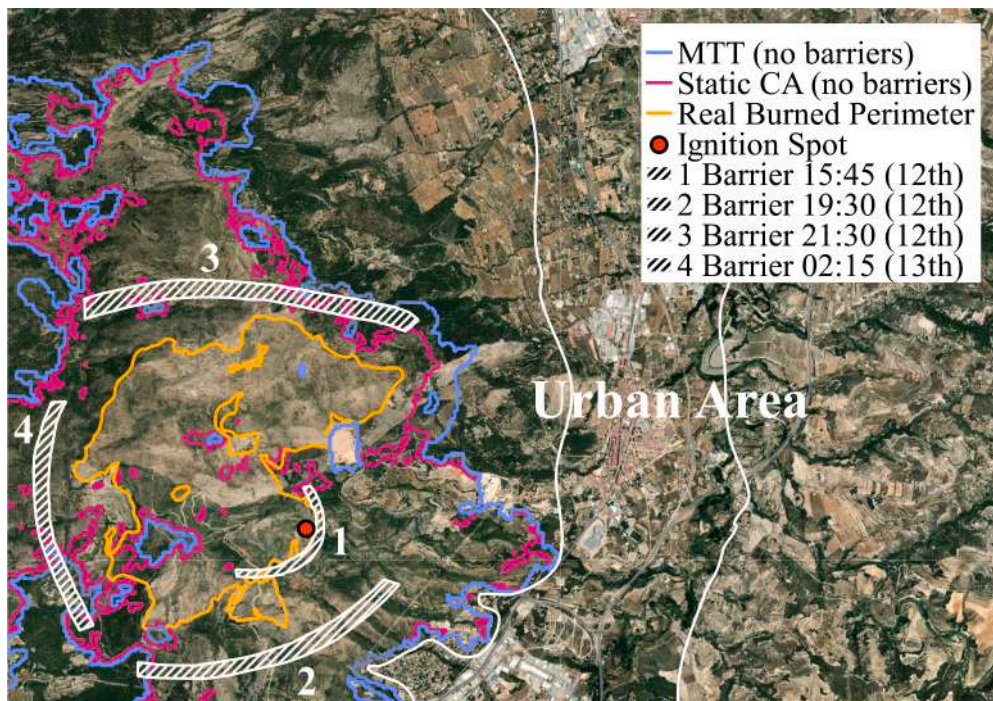- The complete scenario in which both moisture and wind change dynamically.

Fig. 10. Comparison between the FlamMap MTT fire growth algorithm and the static CA fire perimeter predictions. At the left and bottom borders, both MTT and CA dynamics encounter the system boundaries. Barriers are not considered. Simulation time finishes at 6:30 am, 13th (17 h).
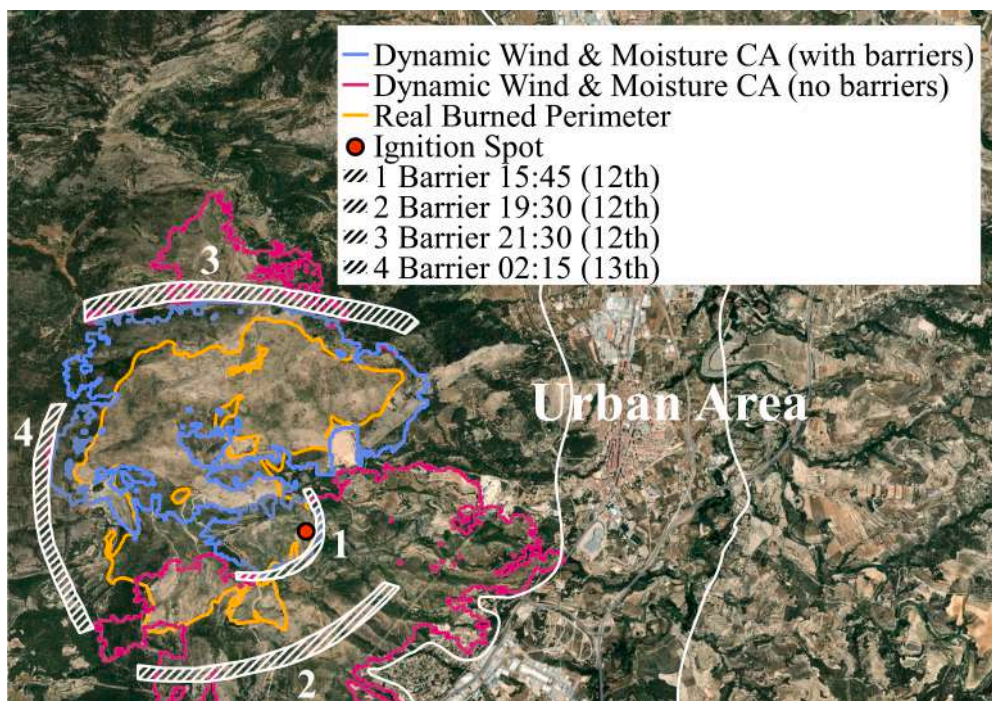
Fig. 13. Results of the CA with dynamic wind and moisture, using eight pre-computed scenarios, without and with barriers. At the bottom border, CA dynamics encounter the system boundary. Simulation time finishes at 6:30 am, 13th (17 h).

Table 1

Model results are presented as the final burned area and percentages of over and under-burned final estimations. First, the static models, where the fire growth model does not consider varying environmental conditions (MTT and equivalent CA). Then, dynamic models without barriers and with dynamic suppression barriers. Results are given per attribute as a range comprising the 95% variability confidence interval of the canopy structure regression model. Each model is presented for 14 h and 17 h completion simulated times.

	Model	Real time (h)	Burned Area $[,]_{95\%}$(Ha)	Over - burned $[,]_{95\%}$(%)	Under - burned $[,]_{95\%}$(%)
Static	MTT	14	[2091,1931]	[287,258]	[3,4]
		17	[2478,2346]	[356,332]	[2,2]
	CA	14	[1711,1582]	[223,202]	[10,12]
		17	[2079,1951]	[286,264]	[5,6]
Dynamic	Wind	14	[1685,1290]	[226,161]	[17,24]
		17	[2166,1625]	[312,215]	[15,17]
	Moisture	14	[1079,975]	[145,127]	[47,48]
		17	[1394,1304]	[194,179]	[38,39]
	Wind &	14	[1121,929]	[133,105]	[28,34]
	Moisture	17	[1546,1325]	[201,160]	[17,17]
Dynamic & Barriers	Wind	14	[695,516]	[56,30]	[28,36]
		17	[839,640]	[81,46]	[27,28]
	Moisture	14	[441,376]	[39,29]	[58,60]
		17	[546,496]	[49,42]	[49,51]
	Wind &	14	[482,370]	[27,13]	[39,46]
	Moisture	17	[640,584]	[46,36]	[29,29]

Table 1 displays the results for static and dynamic fire behavior models, also with the implementation of dynamic suppression barriers.

The predicted burned area intervals indicate that the final perimeter is smaller as the canopy model becomes denser and more extensive. This aligns with the understanding that mid-flame wind magnitude, which enhances fire spread rate, diminishes in the presence of more comprehensive and denser vertical fuel distributions (Andrews, 2012).

MTT and the CA show equivalence in a scenario with no dynamic environment, but some discrepancies are present, as shown in Fig. 10. In the CA, fire spreads in discrete directions and distances, leading to an incremental expansion of the burned surface. A burning cell ignites its neighbors when the fire reaches the exact distance separating them. Consequently, the fire front may traverse neighboring cells during several CA iterations without setting them ablaze. This effect propagates throughout the simulation, resulting in larger predicted MTT burned areas.

Despite identifying these differences, both static approaches display significant over-predicted percentages ranging from 202% to 356%. Without barriers, the CA only significantly improves when moisture and wind are dynamically updated based on weather data, as observed in Fig. 11. In this dynamic model, over-burned percentages indicate that fire growth follows a more realistic pace, reproducing the day-night weather cycle. While limiting wildfire growth, the over-predicted percentages still vary broadly from 105% to 201%. Moreover, considering dynamic winds, canopy uncertainty substantially impacts this result. On the one hand, for thinner and lower canopy layers, the over-predicted burned percentages share values as high as those from the static models (226% to 312%). On the other hand, for denser and taller canopy structures, the over-predicted percentages diminish to 161% and 215%, improving the static results. Regarding the under-burned percentages, these suggest that all three dynamic models fail to burn large areas, up to 38%-48% with dynamic moisture. This last observation states that dynamic moisture has a great sensibility to the day-night cycle, strongly reducing the fire rate of spread.

With the presence of barriers in the dynamic models, over-burned percentages decrease to 81% in the worst-case and 13% in the best result. The inclusion of barriers still validates the model that incorporates both dynamic wind and moisture as the one predicting less excess in the burned surface. Similarly to the scenario without barriers, the model that only considers dynamic winds shows higher dispersion due to uncertainty in the canopy structure. Focusing on under-predicted burned percentages, deploying dynamic barriers in the models increases this magnitude in all cases. Again, when only dynamic moisture is considered for computing time-varying fire behavior scenarios, it
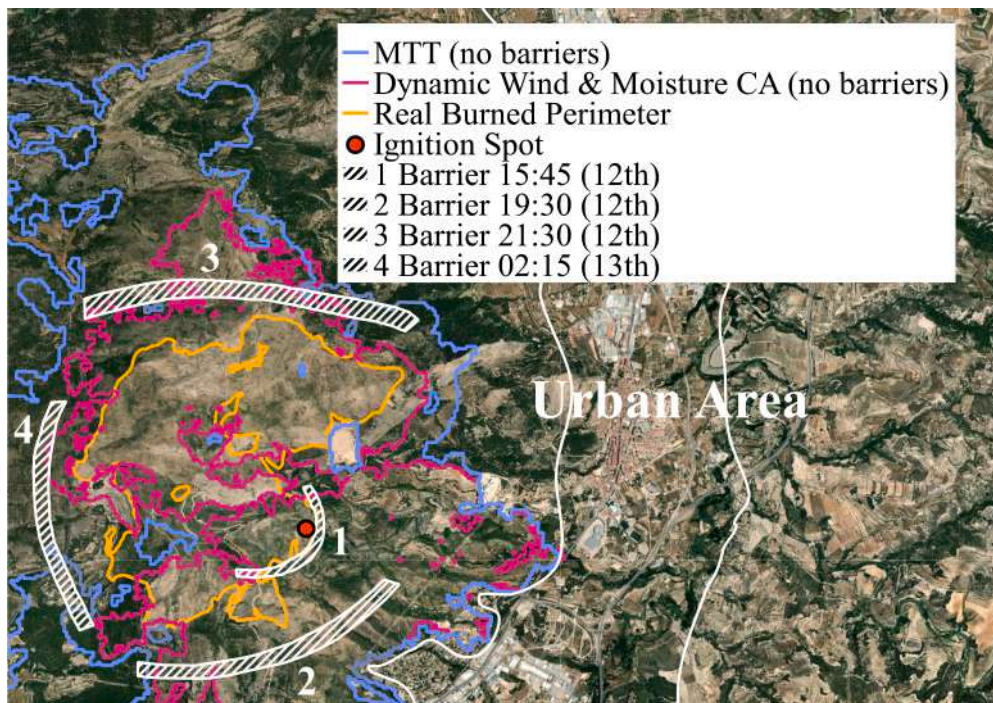
Fig. 11. Comparison between the FlamMap MTT fire growth algorithm and the CA with dynamic wind and moisture, using eight pre-computed scenarios fire perimeter predictions. At the left and bottom borders, both MTT and CA dynamics encounter the system boundaries. Barriers are not considered. Simulation time finishes at 6:30 am, 13th (17 h).

displays the larger under-predicted burned percentages (49% to 60%). The situation in which only wind changes dynamically presents the lower under-predicted burned percentages. Winds consistently strengthen fire spread. Despite the results consistently exhibiting larger variability due to canopy in this setting, under-predicted burnt percentages have smaller magnitudes.

The relative change in over-predicted burned percentages is more remarkable than the increment of under-predicted percentages, along with the inclusion of barriers. Furthermore, over-predicted burned percentages intrinsically feature more variability than under-predicted percentages. It suggests the possibility of misclassified fuel models in the area, leading to lower fire spread rates. If variability in forest structure would introduce variability in over-burned percentages, disturbances should similarly impact under-burned percentages.
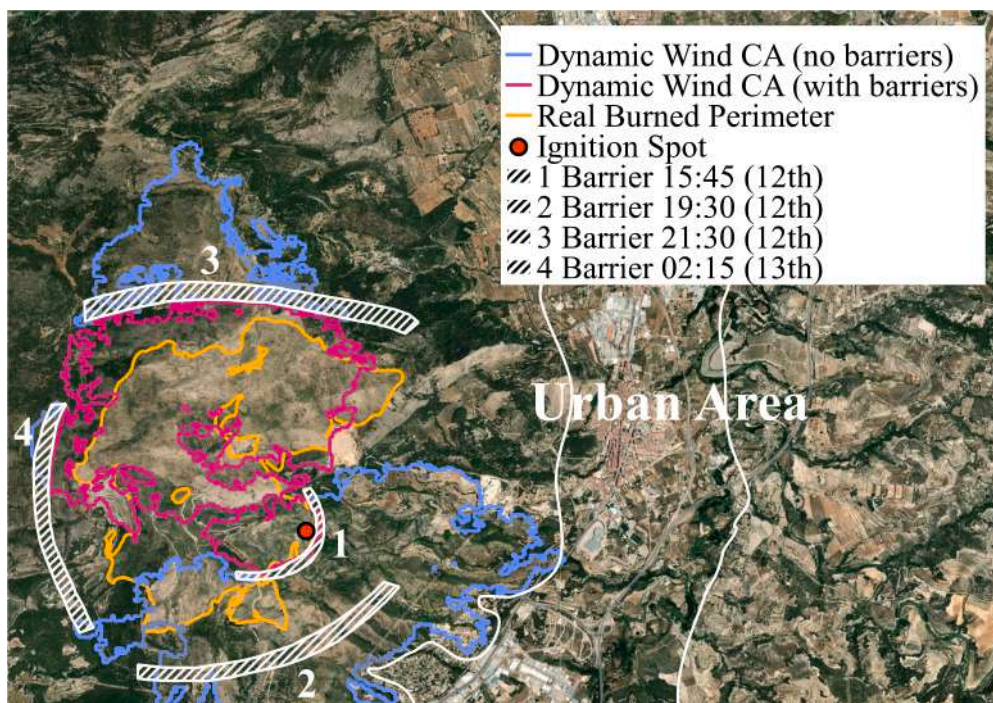

Fig. 12. Results of the CA with dynamic wind, using seven pre-computed scenarios, without and with barriers. At the bottom border, CA dynamics encounter the system boundary. Simulation time finishes at 2:30 am, 13th (14 h).

Fig. 12 and Fig. 13 present results of the different dynamic models considering as canopy layers the median H, CBH, and CBD values in the scenarios with and without barriers. Without barriers, these results illustrate that fire spread extends beyond the recorded final fire perimeter in the southeast area. It has to be highlighted that the recorded ignition point lies approximately at the final burned surface perimeter. However, the local elliptical spread rate always predicts a non-negligible fire spread in the opposite direction to the actual final burned surface, regardless of wind conditions. The role of suppression tactics is inferred to cause this deviation between models and the actual final burned perimeter. In the models, while barriers 2 and 3 prevent the fire front from expanding beyond its borders, barrier 1 significantly affects the southeast direction of fire spread. Thus, initial fire attack efforts can explain how fire did not spread in that direction where urban areas sprawl. Barrier 3 notably prevents northward spread in all considered scenarios, as inferred from media reports and the final fire perimeter. As commented before, the fire initially spread upwards, turned west, and finally northeast. The model predicts that behavior reasonably well.

Simulation completion times have different impacts on the final fire perimeter. Dynamic wind and moisture preconditioning yields better similarity to the final burned surface when the model iterates larger real-time intervals (17 h vs. 14 h). Conversely, when only wind is computed dynamically, the effect is the opposite. This result agrees with the fact that moisture conditioning mirrors the day-night cycle. Temperature and relative humidity fluctuate more than wind during this period. Moisture variability constrains fire spread rates. Fig. 12 and Fig. 13 highlight these aspects, with dynamic wind displaying a simulation of 14 real-time h and with moisture preconditioning simulating 17 real-time h.

Computation times are registered to be about the order of minutes, validating this modeling framework as a viable operational tool. Experiments were run on a laptop computer with 16 Gb RAM, an Intel I7 processor, and using a single core.

As discussed in this work, operational wildfire models inherently display inaccuracies. Although this study enhances our comprehension of these phenomena, several key limitations need consideration:

- The modeling of input uncertainties can be extended to other relevant variables, such as wind data and fuel mapping. A more comprehensive sensitivity analysis yields more accurate fire-related confidence intervals.
- The fire growth model lacks environmental criteria such that under certain circumstances, fuels can be assumed to be virtually exhausted (a natural stopping criterion for fire growth).
- Fire-Wind Feedback: The model does not account for the bidirectional relationship between fire and wind. This feedback mechanism plays a pivotal role in the chaotic dynamics of specific fires (Kartsios et al., 2021), and future operational fire models should endeavor to incorporate this phenomenon.
- This study underscores the significance of human factors as relevant variables in fire dynamics. Other human-related activities can impact fire behavior at the Wildland-Urban-Interface (Wahlqvist et al., 2021; Vacca et al., 2020; Katzilieris et al., 2022) or even cultural, political, and socioeconomic factors can influence fire risks (Cantizano et al., 2022; De Diego et al., 2023).

5. Discussion and conclusions

Global warming has led to an increase in the size and frequency of wildfires. This serious problem requires accurate models that can predict their behavior and help prevent them. These models must consider both micro-level factors, such as physical models of fire propagation and local terrain features, and macro-level factors, such as wind patterns and firefighters' interventions. By providing scientifically rigorous predictions, these models can assist risk management processes and policymakers strategies on the mitigation and impact of wildfires.

Our work introduces a new methodology that utilizes advanced operational models incorporating physical fire propagation models, LiDAR databases, and real-time measurements of wind and humidity. This method enables us to identify the key factors that contribute to the spread of wildfires, such as detailed information on the intervention of firefighters and the local meteorological conditions.

It has traditionally been considered safer to over-predict the spread of fires. However, we emphasize that models should be calibrated based on specific canopy structures and meteorological conditions rather than relying on summary indicators that do not capture the underlying dynamics.

Modeling the spread of fire presents several challenges, including limitations and the inability to capture certain phenomena within current frameworks adequately. Determining simulation end times and assessing the extent of firefighters' ability to influence wildfires are examples of the complexities that require attention.

In this study, we found that wind is the primary factor shaping the pattern of a wildfire, followed by moisture preconditioning. Additionally, the structure of the canopy significantly impacts the predictions, particularly the extent of over-burned areas, highlighting its role in influencing the speed of fire spread over extended periods. While fuel models are classified according to a predefined scheme, there is always uncertainty and variability within this layer of information. However, the misclassification of fuel models is evident through the robustness of the under-burned percentage, indicating that certain areas' fuel models may have needed to be accurately assigned.

The lack of precise information can be mitigated by incorporating uncertainties derived from public data repositories. By accounting for these uncertainties, fire behavior models can better capture the inherent variability and improve their predictive capabilities. While this study does not aim to provide a comprehensive sensitivity analysis, it underscores the importance of considering and incorporating uncertainties to enhance the accuracy and reliability of fire spread modeling in similar contexts.

Finally, we want to emphasize the versatility and accuracy of CA as a fire growth model. The static CA case deviates from the MTT algorithm, a well-accepted approach to wildfire modeling. Nevertheless, differences due to spatial discretization fade away in complex scenarios encountered in Nature. Many works indeed aim to force CA to resemble more FARSITE and MTT. However, results claim this is unnecessary, at least within the scope of this studied wildfire.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Spanish public data repositories

This Appendix explains the three public forestry repositories employed in this work in more detail. Fig. A.1 represents the type of available data from the different repositories.
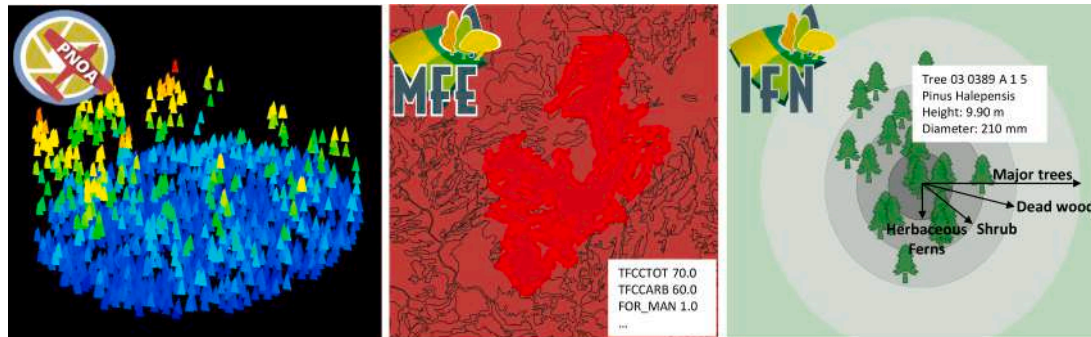

Fig. A.1. Available data from public repositories. Left: LiDAR data is visualized using (McGaughey, 2022). Center: Spanish Forest Map is displayed using (QGIS Development Team, 2021). Right: visual representation of the Spanish National Inventory is taken from the online visualization tool (Vega-Gorgojo et al., 2022).

A.1. Spanish National Orthophoto Program (PNOA)

LiDAR is a remote sensing technique that captures detailed information about objects and vegetation in three-dimensional space. It operates by emitting laser pulses from a source and measuring the time it takes for the reflected pulses to return. This enables the creation of highly accurate and detailed maps of the surveyed area. The output of a LiDAR survey is a collection of point clouds. This technique is particularly well-suited for forestry inventories, as Airborne Laser Scanning (ALS) allows for efficiently characterizing large vegetated areas. The amount of countries that have invested in their public programs to conduct LiDAR surveys is limited, such as Canada, Switzerland, Finland, and Spain. In Spain, the national LiDAR dataset is part of the PNOA project, which aims to cover the entire country in a six-year cycle. The first coverage of the project was conducted between 2009 and 2015, while the second coverage began in 2015 and was completed in 2021. Currently, the third coverage is underway, starting in 2023. These successive LiDAR coverages provide valuable data for various applications, including forestry analysis, land management, and environmental monitoring.

A.2. Spanish Forest Map (MFE)

Fuel model mapping is the basis for surface fire spread modeling. An up-to-date and comprehensive forest cartography mapping, which encompasses vital information about soil utilization, vegetation status, and dominant forest species, is crucial. In Spain, the MFE serves as the national baseline for forest topography, offering a vector representation of forestry ecosystems across the country's landscape, partitioned into homogeneous subdomains. This information is available from the *Ministry for the Ecological Transition and the Demographic Challenge* web in a vector file ready to be processed in GIS software. As part of a periodically updated survey, the versions valuable to this study are the MFE50 (1:50000 scale) and MFE25 (1:25000 scale) maps. The MFE50 was developed between 1998 and 2007, while the MFE25 project commenced in 2007 and is currently ongoing.

The databases describe the forest stands encompassing various fields relating to their ecological and structural features. For tree forest usage, MFE50 considers three distinct species, each with its corresponding stage of development (reforested, coppice, latizal, and fustal). Additionally, it includes the occupation percentage, representing the species' coverage about the total number of trees, as well as the fraction of the total area covered by each of the species. The most recent forest map version MFE25 incorporates the Rothermel Fuel classification as an attribute associated with every tile.

A.3. Spanish National Forest Inventory (IFN)

A forest inventory involves systematically collecting forest data and information from precisely defined plot stands for assessment or analysis. The spatial distribution of these plots ideally aims to statistically represent the variables determining the forest structure in the surveyed area.

Unlike the MFE, which relies primarily on remotely acquired data, forest inventories are exclusively recorded on-site by specialized personnel equipped with the necessary tools and instruments. There is no standardized methodology for characterizing forest geometrical structures using remote sensing techniques.

In Spain, the IFN project carries out a comprehensive update of the inventory periodically every ten years, starting in 1966. The *Fourth National Forest Inventory* (IFN4) is the current iteration.

References

Albini, Frank A. 1976. Estimating wildfire behavior and effects. Gen. Tech. Rep. INT-GTR-30. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 92 p. https://www.fs.usda.gov/research/treesearch/29574.

Alexandridis, A., Russo, L., Vakalis, D., Bafas, G.V., Siettos, C.I., 2011. Wildland fire spread modelling using cellular automata: evolution in large-scale spatially heterogeneous environments under fire suppression tactics. Int. J. Wildland Fire 20, 633. https://doi.org/10.1071/WF09119.

Andrews, P.L., 2012. Modeling wind adjustment factor and midflame wind speed for Rothermel's surface fire spread model. Technical Report. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research, Station. https://doi.org/10.2737/RMRS-GTR-266.

Arca, B., Duce, P., Laconi, M., Pellizzaro, G., Salis, M., Spano, D., Arca, B., Duce, P., Laconi, M., Pellizzaro, G., Salis, M., Spano, D., 2007. Evaluation of FARSITE simulator in Mediterranean maquis. Int. J. Wildland Fire 16, 563–572. https://doi.org/10.1071/WF06070.

Arumäe, T., Lang, M., 2018. Estimation of canopy cover in dense mixed-species forests using airborne lidar data. Eur. J. Remote Sens. 51, 132–141. https://doi.org/10.1080/22797254.2017.1411169.

Benali, A., Ervilha, A.R., Sá, A.C.L., Fernandes, P.M., Pinto, R.M.S., Trigo, R.M., Pereira, J.M.C., 2016. Deciphering the impact of uncertainty on the accuracy of large wildfire spread simulations. Sci. Total Environ. 569–570, 73–85. https://doi.org/10.1016/j.scitotenv.2016.06.112.

Bowman, D.M.J.S., Kolden, C.A., Abatzoglou, J.T., Johnston, F.H., van der Werf, G.R., Flannigan, M., 2020. Vegetation fires in the Anthropocene. Nat. Rev. Earth Environ. 1, 500–515. https://doi.org/10.1038/s43017-020-0085-3.

Canadian Forest Service Publications, 1978. Canadian forest fire weather index tables. Forestry Technical Report 25. https://cfs.nrcan.gc.ca/publications?id=24013.

Cantizano, A., Caro, R., Fernández, M., Ayala, P., 2022. Human factors in the model of urban fire spread in Madrid (Spain) focused on the poor population. Sustainability 14, 4486. https://doi.org/10.3390/su14084486.

Considine, G.D., 2006. National Fire Danger Rating System (NFDRS). John Wiley & Sons, Inc, Hoboken, NJ, USA. https://doi.org/10.1002/0471743984.vse8649.

Crimaldi, M., Lama, G.. Impacts of riparian plants biomass assessed by uav-acquired multispectral images on the hydrodynamics of vegetated streams. https://doi.org/10.5071/29thEUBCE2021-4AV.3.6.

Cruz, M.G., Alexander, M.E., 2013. Uncertainty associated with model predictions of surface and crown fire rates of spread. Environ. Model. Software 47, 16–28. https://doi.org/10.1016/j.envsoft.2013.04.004.

De Diego, J., Fernández, M., Rúa, A., Kline, J.D., 2023. Examining socioeconomic factors associated with wildfire occurrence and burned area in Galicia (Spain) using spatial and temporal data. Fire. Ecology 19, 18. https://doi.org/10.1186/s42408-023-00173-8.

Errico, A., Lama, G., Francalanci, S., Chirico, G., Solari, L., Preti, F.. Validation of global flow resistance models in two experimental drainage channels covered by Phragmites australis (common reed). https://doi.org/10.3850/38WC092019-1215.

Fidalgo-González, L.A., Arellano-Pérez, S., Álvarez González, J.G., Castedo-Dorado, F., Ruiz-González, A.D., González-Ferreiro, E., 2019. Estimation of the vertical distribution of the fine canopy fuel in Pinus sylvestris stands using low density LiDAR data. Revista de Teledetección. https://doi.org/10.4995/raet.2019.11241.

Finney, M.A., 1998. FARSITE: Fire Area Simulator-model development and evaluation. Res. Pap. RMRS-RP-4, Revised 2004. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 47 p. 4. https://doi.org/10.2737/RMRS-RP-4.

Finney, M.A., 2002. Fire growth using minimum travel time methods. Can. J. For. Res. 32, 1420–1424. https://doi.org/10.1139/x02-068.

Finney, M.A., 2006. An Overview of FlamMap Fire Modeling Capabilities. In: Andrews, Patricia L.; Butler, Bret W., comps. 2006. Fuels Management-How to Measure Success: Conference Proceedings. 28–30 March 2006; Portland, OR. Proceedings RMRS-P-41. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 213–220 041. http://www.fs.usda.gov/treesearch/pubs/25948.

Finney, M.A., Grenfell, I.C., McHugh, C.W., Seli, R.C., Trethewey, D., Stratton, R.D., Brittain, S., 2011. A method for ensemble wildland fire simulation. Environ. Model. Assess. 16, 153–167. https://doi.org/10.1007/s10666-010-9241-3.

Forthofer, J.M., 2007. Modeling Wind in Complex Terrain for use in Fire Spread Prediction. Master of Science. Fort Collins, CO: Colorado State University, 123 p. https://www.frames.gov/catalog/8132.

González-Ferreiro, E., Arellano-Pérez, S., Castedo-Dorado, F., Hevia, A., Vega, J.A., Vega-Nieva, D., Álvarez González, J.G., Ruiz-González, A.D., 2017. Modelling the vertical distribution of canopy fuel load using national forest inventory and low-density airbone laser scanning data. PLOS ONE 12 e0176114. https://doi.org/10.1371/journal.pone.0176114.

Guo, M., Yao, Q., Suo, H., Xu, X., Li, J., He, H., Yin, S., Li, J., 2023. The importance degree of weather elements in driving wildfire occurrence in mainland China. Ecol. Ind. 148, 110152. https://doi.org/10.1016/j.ecolind.2023.110152.

Haines, D., 1988. A lower atmospheric severity index for wildland fires. National Weather Digest 13, 23–27. https://www.frames.gov/catalog/11664.

Kartsios, S., Karacostas, T., Pytharoulis, I., Dimitrakopoulos, A.P., 2021. Numerical investigation of atmosphere-fire interactions during high-impact wildland fire events in Greece. Atmos. Res. 247, 1052. https://doi.org/10.1016/j.atmosres.2020.105253.

Katzilieris, K., Vlahogianni, E.I., Wang, H., 2022. Evacuation behavior of affected individuals and households in response to the 2018 attica wildfires: From empirical data to models. Saf. Sci. 153, 105799. https://doi.org/10.1016/j.ssci.2022.105799.

Keetch, John J., Byram, George M., 1968. In: A Drought Index for Forest Fire Control. Res. Pap. SE-38, 35. U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, Asheville, NC. https://www.fs.usda.gov/research/treesearch/40.

González-Olabarria, JR., Piqué, M. Busquets, E. 2019. Cartografia de vegetació per la simulació d'incendis forestals [Vegetation mapping for simulating forest fires]. Servidor PREVINCAT. https://previncat.ctfc.cat/.

Kelly, M., Su, Y., Di Tommaso, S., Fry, D., Collins, B., Stephens, S., Guo, Q., 2017. Impact of Error in Lidar-Derived Canopy Height and Canopy Base Height on Modeled Wildfire Behavior in the Sierra Nevada, California, USA. Remote Sensing 10, 10. https://doi.org/10.3390/rs10010010.

Lama, G., Crimaldi, M. Assessing the role of gap fraction on the leaf area index (lai) estimations of riparian vegetation based on fisheye lenses. https://doi.org/10.5071/29thEUBCE2021-4AV.3.16.

Lama, G.F.C., Crimaldi, M., De Vivo, A., Chirico, G.B., Sarghini, F., 2021. Eco-hydrodynamic characterization of vegetated flows derived by UAV-based imagery. In: 2021 IEEE International Workshop on Metrology for Agriculture and Forestry

(MetroAgriFor). IEEE, Trento-Bolzano, Italy, pp. 273–278. https://doi.org/10.1109/MetroAgriFor52389.2021.9628749.

Lama, G.F.C., Rillo Migliorini Giovannini, M., Errico, A., Mirzaei, S., Chirico, G.B., Preti, F., 2021b. The impacts of Nature Based Solutions (NBS) on vegetated flows' dynamics in urban areas. IEEE, Trento-Bolzano, Italy, pp. 58–63. https://doi.org/10.1109/MetroAgriFor52389.2021.9628438.

Lense, G.H.E., Lämmle, L., Ayer, J.E.B., Lama, G.F.C., Rubira, F.G., Mincato, R.L., 2023. Modeling of Soil Loss by Water Erosion and Its Impacts on the Cantareira System, Brazil. Water 15, 1490. https://doi.org/10.3390/w15081490.

Linn, R., Reisner, J., Colman, J.J., Winterkamp, J., 2002. Studying wildfire behavior using FIRETEC. Int. J. Wildland Fire 11, 233. https://doi.org/10.1071/WF02007.

Liu, N., Lei, J., Gao, W., Chen, H., Xie, X., 2021. Combustion dynamics of large-scale wildfires. Proc. Combust. Inst. 38, 157–198. https://doi.org/10.1016/j.proci.2020.11.006.

McGaughey, R.J.. FUSION/LDV: Software for LIDAR Data Analysis and Visualization. https://forsys.sefs.uw.edu/fusion/fusionlatest.html.

Mell, W., Jenkins, M.A., Gould, J., Cheney, P., 2007. A physics-based approach to modelling grassland fires. Int. J. Wildland Fire 16, 1. https://doi.org/10.1071/WF06002.

Ministry for the Ecological Transition and the Demographic Challenge (MITECO), 2005. Spanish Forestry Mapping (MFE). https://www.miteco.gob.es/en/cartografia-y-sig/ide/descargas/biodiversidad/mfe_comunidadvalenciana.html.

Ministry for the Ecological Transition and the Demographic Challenge (MITECO), 2012. General Statistics on Forest Fires (EGIF). Technical Report. https://www.miteco.gob.es/en/biodiversidad/temas/incendios-forestales/estadisticas-datos.html.

Ministry for the Ecological Transition and the Demographic Challenge (MITECO), 1997. Spanish National Forest Inventory (IFN). https://www.miteco.gob.es/en/biodiversidad/servicios/banco-datos-naturaleza/informacion-disponible/ifn3_base_datos_1_25.html.

Mohammad, L., Bandyopadhyay, J., Sk, R., Mondal, I., Nguyen, T.T., Lama, G.F.C., Anh, D.T., 2023. Estimation of agricultural burned affected area using NDVI and dNBR satellite-based empirical models. J. Environ. Manage. 343, 118226. https://doi.org/10.1016/j.jenvman.2023.118226.

Montero, G., Ruiz-Peinado, R., 2006. Producción de biomasa y fijación de CO2 por los bosques españoles [Biomass production and CO2 fixation in Spanish forests]. ININA - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria. OCLC: 1104440771.ISBN: 84-7498-512-9 .

Mutthulakshmi, K., Wee, M.R.E., Wong, Y.C.K., Lai, J.W., Koh, J.M., Acharya, U.R., Cheong, K.H., 2020. Simulating forest fire spread and fire-fighting using cellular automata. Chin. J. Phys. 65, 642–650. https://doi.org/10.1016/j.cjph.2020.04.001.

Myoung, B., Kim, S., Nghiem, S., Jia, S., Whitney, K., Kafatos, M., 2018. Estimating Live Fuel Moisture from MODIS Satellite Data for Wildfire Danger Assessment in Southern California USA. Remote Sens. 10, 87. https://doi.org/10.3390/rs10010087.

Nelson Jr., R.M., 2000. Prediction of diurnal change in 10-h fuel stick moisture content. Can. J. For. Res. 30, 1071–1087. https://doi.org/10.1139/x00-032.

Oliveira, S., Rocha, J., Sá, A., 2021. Wildfire risk modeling. Curr. Opin. Environ. Sci. Health 23, 100274. https://doi.org/10.1016/j.coesh.2021.100274.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830. https://doi.org/10.5555/1953048.2078195.

Pimont, F., Dupuy, J.L., Linn, R.R., Dupont, S., 2009. Validation of FIRETEC wind-flows over a canopy and a fuel-break. Int. J. Wildland Fire 18, 775–790. https://doi.org/10.1071/WF07130.

Pirone, D., Cimorelli, L., Del Giudice, G., Pianese, D., 2023. Short-term rainfall forecasting using cumulative precipitation fields from station data: a probabilistic machine learning approach. J. Hydrol. 617, 128949. https://doi.org/10.1016/j.jhydrol.2022.128949.

Purnomo, D.M.J., Bonner, M., Moafi, S., Rein, G., 2021. Using cellular automata to simulate field-scale flaming and smouldering wildfires in tropical peatlands. Proc. Combust. Inst. 38, 5119–5127. https://doi.org/10.1016/j.proci.2020.08.052.

QGIS Development Team, 2021. QGIS Geographic Information System. Open Source Geospatial Foundation. http://qgis.org.

rapidlasso GmbH, LAStools, Efficient LiDAR Processing Software (version 220107, open license), obtained from http://rapidlasso.com/LAStools.

Ray, R., Das, A., Hasan, M.S.U., Aldrees, A., Islam, S., Khan, M.A., Lama, G.F.C., 2023. Quantitative Analysis of Land Use and Land Cover Dynamics using Geoinformatics Techniques: A Case Study on Kolkata Metropolitan Development Authority (KMDA) in West Bengal, India. Remote Sens. 15, 959. https://doi.org/10.3390/rs15040959.

Reeves, M.C., Ryan, K.C., Rollins, M.G., Thompson, T.G., 2009. Spatial fuel data products of the LANDFIRE Project. Int. J. Wildland Fire 18, 250. https://doi.org/10.1071/WF08086.

Rothermel, R.C., 1991. Predicting behavior and size of crown fires in the northern Rocky Mountains. Res. Pap. INT-438. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 46 p. 438. https://doi.org/10.2737/INT-RP-438.

Runfola, D., Contributors, C., Rogers, L., Habib, J., Horn, S., Murphy, S., Miller, D., Day, H., Troup, L., Fornatora, D., Spage, N., Pupkiewicz, K., Roth, M., Rivera, C., Altman, C., Schruer, I., McLaughlin, T., Biddle, R., Ritchey, R., Topness, E., Turner, 2020. geoBoundaries: A global database of political administrative boundaries. In: Updike, S., Buckman, H., Simpson, N., Lin, J., Anderson, A., Baier, H., Crittenden, M., Dowker, E., Fuhrig, S., Goodman, S., Grimsley, G., Layko, R., Melville, G., Mulder, M., Oberman, R., Panganiban, J., Peck, A., Seitz, L., Shea, S.,

Slevin, H., Yougerman, R., Hobbs, L. (Eds.), Plos one 15, e0231866. https://doi.org/10.1371/journal.pone.0231866.

Sakia, R.M., 1992. The box-cox transformation technique: a review. Journal of the Royal Statistical Society Series D: The Statistician 41, 169–178. https://doi.org/10.2307/2348250.

Seydi, S.T., Hasanlou, M., Chanussot, J., 2022. Burnt-Net: Wildfire burned area mapping with single post-fire Sentinel-2 data and deep learning morphological neural network. Ecol. Ind. 140, 108999. https://doi.org/10.1016/j.ecolind.2022.108999.

Shakila, B.G., óscar, L., Adrián, U., 2012. Así estamos contando el incendio en Mariola [This is how we are narrating the fire in Mariola]. Pagina 66. https://pagina66.com/art/55681/asi-estamos-contando-el-incendio-en-mariola.

Soriano, J.L., Botella, M.A., 2015. Sistema Integrado de Gestión de Incendios Forestales (SIGIF), Informes Post Incendio Compendio Anual 2012–2013 [Integrated Wildland Fire Management System (SIGIF), Post-Fire Reports Annual Compendium 2012–2013]. Technical Report. https://prevencionincendiosgva.es/Inicio.

Spanish National Geographic Institute (IGN). PNOA LiDAR First Survey. https://centrodescargas.cnig.es/CentroDescargas/locale?request_locale=en.

Stocks, B.J., Alexander, M.E., Wotton, B.M., Stefner, C.N., Flannigan, M.D., Taylor, S.W., Lavoie, N., Mason, J.A., Hartley, G.R., Maffey, M.E., Dalrymple, G.N., Blake, T.W., Cruz, M.G., Lanoville, R.A., 2004. Crown fire behaviour in a northern jack pine black spruce forest. Can. J. For. Res. 34, 1548–1560. https://doi.org/10.1139/x04-054.

Sánchez García, S., García, M., Velasco, A., Canga, E., 2019. Generación de cartografía de modelos de combustible a partir de datos LiDAR: herramienta flexible, actualizable y escalable [Generation of fuel model mapping from LiDAR data: flexible, updatable and scalable tool]. Revista TRESEME44. 44, 24–27. https://asmadera.com/wp-content/uploads/2014/11/TRESEME44.pdf.

Sullivan, A.L., 2009a. Wildland surface fire spread modelling, 1990–2007. 1: Physical and quasi-physical models. International Journal of Wildland Fire 18, 349. https://doi.org/10.1071/WF06143.

Sullivan, A.L., 2009b. Wildland surface fire spread modelling, 1990–2007. 2: Empirical and quasi-empirical models. International Journal of Wildland Fire 18, 369. https://doi.org/10.1071/WF06142.

Sullivan, A.L., 2009c. Wildland surface fire spread modelling, 1990–2007. 3: Simulation and mathematical analogue models. International Journal of Wildland Fire 18, 387. https://doi.org/10.1071/WF06144.

Trucchia, A., D'Andrea, M., Baghino, F., Fiorucci, P., Ferraris, L., Negro, D., Gollini, A., Severino, M., 2020. PROPAGATOR: An Operational Cellular-Automata Based Wildfire Simulator. Fire 3, 26. https://doi.org/10.3390/fire3030026.

Sando, Rodney W.; Wick, Charles H. 1972. A method of evaluating crown fuels in forest stands. Research Paper NC-84. St. Paul, MN: U.S. Dept. of Agriculture, Forest Service, North Central Forest Experiment Station. https://www.fs.usda.gov/research/treesearch/10605.

Scott, Joe H.; Burgan, Robert E. 2005. Standard fire behavior fuel models: a comprehensive set for use with Rothermel's surface fire spread model. Gen. Tech.

Rep. RMRS-GTR-153. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 72 p. https://doi.org/10.2737/RMRS-GTR-153.

Tymstra, C., Bryce, R., Wotton, B., Taylor, S., Armitage, O., et al., 2010. Development and structure of Prometheus: the Canadian wildland fire growth simulation model. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Information Report NOR-X-417. (Edmonton, AB). https://cfs.nrcan.gc.ca/publications?id=31775.

Vacca, P., Caballero, D., Pastor, E., Planas, E., 2020. WUI fire risk mitigation in europe: A performance-based design approach at home-owner level. J. Saf. Sci. Resilience 1, 97–105. https://doi.org/10.1016/j.jnlssr.2020.08.001.

Vega-Gorgojo, G., Ordóñez, C., Giménez-García, J.M., Bravo, F., 2022. Explorando datos abiertos forestales masivos con un navegador web [Exploring massive open forest data with a web browser]. Ecosistemas 31, 2452. https://doi.org/10.7818/ECOS.2452. URL: https://www.revistaecosistemas.net/index.php/ecosistemas/article/view/2452.

Wagenbrenner, N.S., Forthofer, J.M., Lamb, B.K., Shannon, K.S., Butler, B.W., 2016. Downscaling surface wind predictions from numerical weather prediction models in complex terrain with WindNinja. Atmos. Chem. Phys. 16, 5229–5241. https://doi.org/10.5194/acp-16-5229-2016.

Wagner, C.E.V., 1977. Conditions for the start and spread of crown fire. Can. J. For. Res. 7, 23–34. https://doi.org/10.1139/x77-004.

Wahlqvist, J., Ronchi, E., Gwynne, S.M., Kinateder, M., Rein, G., Mitchell, H., Bénichou, N., Ma, C., Kimball, A., Kuligowski, E., 2021. The simulation of wildland-urban interface fire evacuation: The wui-nity platform. Saf. Sci. 136, 105145. https://doi.org/10.1016/j.ssci.2020.105145.

Weise, D.R., Koo, E., Zhou, X., Mahalingam, S., Morandini, F., Balbi, J.H., 2016. Fire spread in chaparral - a comparison of laboratory data and model predictions in burning live fuels. Int. J. Wildland Fire 25, 980. https://doi.org/10.1071/WF15177.

White, J.C., Tompalski, P., Vastaranta, M., Wulder, M., Saarinen, N., Stepper, C., Coops, N., 2017. A model development and application guide for generating an enhanced forest inventory using airborne laser scanning data and an area-based approach. CWFC Information Report FI-X-018, 38. https://cfs.nrcan.gc.ca/publications?id=38945.

Wegrzynski, W., Lipecki, T., 2018. Wind and Fire Coupled Modelling-Part I: Literature Review. Fire Technol 54, 1405–1442. https://doi.org/10.1007/s10694-018-0748-5.

Yousefi, S., Avand, M., Yariyan, P., Pourghasemi, H.R., Keesstra, S., Tavangar, S., Tabibian, S., 2020. A novel GIS-based ensemble technique for rangeland downward trend mapping as an ecological indicator change. Ecol. Ind. 117, 106591. https://doi.org/10.1016/j.ecolind.2020.106591.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. Roy. Stat. Soc. Ser. B: Stat. Methodol. 67, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.