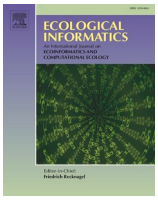
Ecological Informatics 80 (2024) 102456

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



# An intelligent cellular automaton scheme for modelling forest fires☆

Joan Boters-Pitarch [a], María Teresa Signes-Pont [a,*], Julian Szymański [b], Higinio Mora-Mora [a]

[a] Department of Computer Science and Technology, University of Alicante, Alicante, Spain
[b] Faculty of ETI, Gdańsk University of Technology, Gdańsk, Poland

ARTICLE INFO

*Keywords:*
Forest fires
Spread models
Intelligent architecture
Neighbourhood relationship

ABSTRACT

Forest fires have devastating consequences for the environment, the economy and human lives. Understanding their dynamics is therefore crucial for planning the resources allocated to combat them effectively. In a world where the incidence of such phenomena is increasing every year, the demand for efficient and accurate computational models is becoming increasingly necessary. In this study, we perform a revision of an initial proposal which consists of a two-dimensional propagation model based on cellular automata (2D-CA), which aims to understand the dynamics of these phenomena. We identify the key theoretical weaknesses and propose improvements to address these limitations. We also assess the effectiveness and accuracy of the model by evaluating improvements using real forest fire data (Beneixama, Alicante 2019). Moreover, as a result of the theoretical modifications performed, we introduce a novel intelligent architecture that seeks to capture relationships between system cells from the data. This new architecture has the ability to advance our understanding of forest fire dynamics, contributing to both the evaluation of existing protocols and more efficient firefighting resource management.

## 1. Introduction

Mathematical models are of great concern in scientific research and essential tools to approach the modelling of natural events, since they allow the simulation and analysis of the behaviour of complex systems. As well-known examples we mention the spread of wildfires (Hernández Encinas et al., 2007a; Hernández Encinas et al., 2007b; Karafyllidis and Thanailakis, 1997), population dynamics (Malthus et al., 1992; Wangersky, 1978), the evolution of infectious and degenerative diseases (Kermark and Mckendrick, 1927; Kuang et al., 2018) or even climate estimation (Edwards, 2011). These models are also an aid for the development of prevention protocols against natural disasters. However, we encounter several challenges in this field that require obtaining accurate and reliable results. One of the most significant problems is the large number of variables that can be involved in the development of a natural event, which makes the model very complex, so powerful and precise mathematical tools must be used to handle the large amount of data. In addition, linear models may not be suitable for these situations because of their simplicity. Another critical challenge is predicting the future of natural events due to their chaotic nature. As an example, small variations in initial conditions can lead to significant changes in the final outcome of a natural event. To overcome this drawback, scientists have developed advanced techniques such as chaos theory (Boccaletti et al., 2000).

Finally, the lack of quality data can make it difficult to process information and estimate numerical values for some variables. Scientists may need to use appropriate methodologies to obtain the best approximations, such as estimating numerical values for unknown variables (Amat et al., 2002; Emmanuel et al., 2021; Saar-Tsechansky and Provost, 2007) or collecting data more accurately. This article will focus on examining wildfires, to understand the causes and consequences that are crucial for the development of effective preventive measures and management strategies. According to (Ganteaume et al., 2013), the causes of forest fires vary and differ between countries. In Europe, particularly in the Mediterranean region, human activities, particularly arson, are the primary cause of fires. Socioeconomic factors, such as unemployment rates and variables associated with agricultural activities, can also influence the initiation of intentional and unintentional fires, depending on the regional context. Environmental factors, including climate, combustible plant mass and orography, have the most significant impact on forest fire ignition, particularly in Mediterranean-type regions, see Fig. 1.

☆ This document is the results of the research project AICO/2021/331 funded by Generalitat Valenciana.
* Corresponding author.
*E-mail address:* teresa@dtic.ua.es (M.T. Signes-Pont).

https://doi.org/10.1016/j.ecoinf.2023.102456
Received 16 September 2023; Received in revised form 27 December 2023; Accepted 30 December 2023
Available online 8 January 2024

Moreover, these events have catastrophic consequences such as soil degradation, erosion, loss of biodiversity, and reduced air quality (Földi and Kuti, 2016). At the same time, the impact of forest fires can be substantial, encompassing both economic effects, such as property damage and reduced tourism in the affected area, and social consequences, such as depopulation and loss of homes.

Due to the characteristics of the system under study, our model is based on cellular automata (CA) (Boters-Pitarch et al., 2023a; Boters-Pitarch et al., 2023b), since this model allows us to consider both the spatial relationship between the elements of the system and the factors that influence the spread of forest fires like the wind, which is integrated as a key factor that determines neighbourhood relations. Moreover, a detailed theoretical analysis will be carried out to identify weaknesses that need to be improved. These improvements will increase the accuracy of the model and provide better prevention and management of forest fires. Although there is a lack of experiments or articles testing the effectiveness of CA or similar models using real wildfire data, it is worth noting that some authors attempted to carry out this task, as demonstrated by previous work (Alexandridis et al., 2008; Cruz et al., 2018; Ntinas et al., 2017; Stojanova et al., 2006). Despite these efforts, it is still not common to observe performance metrics for such models when applied to actual wildfires. This study aims to test the effectiveness of the proposed model and its new improvements using real data collected from the forest fire that occurred in Beneixama (Alicante) in 2019 (*Sistema integrado de gestión de incendios forestales: Informes post-incendio*, 2023). As a result, precision metrics will be obtained, at least for this particular example, and a methodology will be designed that can be extended to other examples and forest fires.

Thus, the purpose of our work is to test the model in a realistic scenario, contrasting the results with real data and being able to find out its predictive capacity for this type of phenomenon. Furthermore, our research aims to contribute to the development of new architectures built on the current model, whose estimates can offer a deeper understanding of the dynamics inherent to complex phenomena, such as the one addressed in this article. This research is structured as follows: After Introduction, Section 2 provides a review of the primary characteristics of the model, it also analyzes its limitations and suggests theoretical modifications to improve it. Section 3 presents an overview of the study design, including the measured variables, data collection methods and other relevant details related to implementation. Section 4 presents the results of our model that are discussed. Section 5 shows the new intelligent scheme which could learn the neighbourhood relationship between cells. Finally, Section 6 provides a summary and some concluding remarks.

## 2. Theoretical framework of our model

To clarify the very objective of this research we must put it in context. As mentioned previously, increasing global warming and intense human activities have contributed significantly to the increase in wildfires in recent years. Of the 30 specific causes of ignition, more than 95% of incidents can be attributed to human activities, see (Yang et al., 2023; Ying et al., 2018). Therefore, any attempt to predict the location of ignition (i.e., the focus of the fire) may be unsuccessful. However, we firmly believe that once the fire has started, models can be designed that describe the dynamics of the phenomenon, that is, describe the space-time evolution of the fire, using climatic variables. The objective of the model is then to understand this dynamic through the use of a 2D-CA, with a neighbourhood relationship which is variable and depends on the number of infected elements, on their location and, above all, on the climatic values.

### 2.1. Current model

We now describe the approach used in our model (Boters-Pitarch et al., 2023a) to address the spread of real wind-driven phenomena, highlighting how it can be used as a Monte Carlo method (Boters-Pitarch et al., 2023b) to estimate the probabilities of our classification problem.

The model follows a general CA framework, in which every cell occupies one of the three states of the Susceptible-Infected-Dead (SID) paradigm at each generation. In summary, the model is based on the dynamics inherent to biological systems, where each cell represents an individual agent and the interactions between them contribute to the overall dynamics of the system. Fig. 3 shows a general iterative scheme of the CA model In this model, the neighbourhood relationship $R_k$ is characterized by its flexibility and probabilistic nature, where expansion occurs only by infected cells. In addition, the direction $\theta$ and power $\rho$ of the wind are used as a guide for the propagation, which facilitates a realistic representation of certain climatic phenomena and allows a more accurate simulation of the behaviour of the system. To account for uncertainties, the Bernoulli random variable $Be(p)$ is used to determine whether a given cell is infected or not. In practice, a random extraction is performed to update each cell, using the probabilities derived from the neighbourhood relationship. This approach has been summarized in Fig. 4 through an example. We start at a given state, $S_k$, of the grid where only the central element is infected and the surrounding elements are susceptible to infection. Then, we calculate the infection probabilities for each of the cells according to $\theta$ and $\rho$ (central grid). Finally, our update criterion $U_k$ based on random draws determines the next $S_k + 1$ state of the grid. One of the most interesting contributions of the model is its ability to determine the probabilities of infection (susceptibility or death) of each cell and generation. This is achieved through the stochastic update criterion, which yields different results for different seeds and allows the model to be used as a Monte Carlo method. By applying the law of large numbers, we can assume that the frequency of the event for each cell in each generation provides an estimate of the probability of that state occurring.

### 2.2. Limitations of the model

An analysis of the weaknesses of the model is crucial to make it scalable in terms of inclusion and manageability of new variables. These limitations are described below:

1. **Dependence on wind**: Although wind is an important variable for the spread of the phenomenon, it is not the only factor that affects the

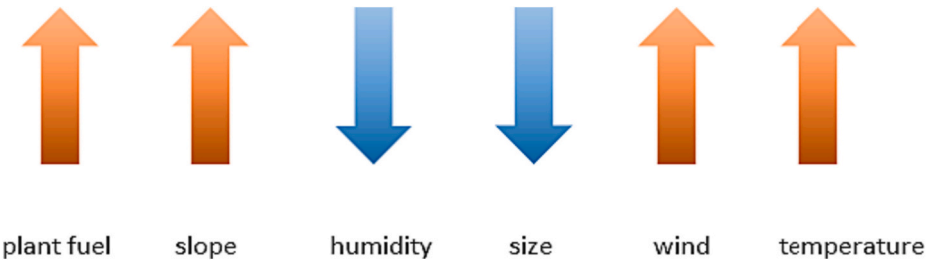



**Fig. 1.** How some environmental factors affect the spread of forest fires.

behaviour of fire in nature. Therefore, the use of wind as the main factor may limit its ability to capture the complexity of the phenomenon under study. To address this drawback, more climate variables are needed. They should be included in the model, although doing so may not be easy since the variables must be adapted to the specific characteristics of each phenomenon, which makes generalization of the model difficult.

2. **Bounded infection capacity**: The infection capacity is restricted for each generation by the model's construction. This may limit the ability of the model to represent realistic fire spread in nature, particularly in situations where fire spreads faster than the current infection capacity can account for.
3. **Limitations of the update rule**: The model uses an update rule to deal with uncertainty, but this function **is not differentiable**, and this could pose challenges if intelligent elements were to be incorporated into the CA. To overcome this limitation, alternative update rules should be explored to improve the ability of the model to represent different situations more accurately.

### 2.3. Proposed improvements

First of all, to minimize the dependence on wind, we transform parameters $C$ and $p_0$ into climate functions that rely on temperature and humidity. Secondly, Definition 3.4 in (Boters-Pitarch et al., 2023a) will be modified to expand the infection capacity by increasing the number of elements in partitions. In other words, we relax the constraints on the infection capacity of the model. This setting seeks to provide more flexibility in capturing the spread of infections by adding a new parameter which is the size of the partitions (i.e. the number of rows and columns potentially affected). Lastly, we explore the benefits of the Gumbel-Softmax function (Jang et al., 2016; Maddison et al., 2016) to make our update criterion differentiable while keeping a similar statistical performance.

#### 2.3.1. Parameters $C$ and $p_0$ as climate functions

We define both, a lower bound $T_{min}$ and a higher bound $T_{max}$ such that the phenomenon under consideration cannot occur for temperature values $t < T_{min}$ and $T_{max}$ is the highest temperature recorded in the geographical area under study. Very low (even negative) temperatures are unlikely to lead to significant spread and impact of forest fires. Thus, we define the domain $D_T = ]T_{min}, T_{max}]$. We also consider the percentage of relative humidity as a parameter, therefore the domain $D_H$ is $]0, 100]$. Therefore, the support is as follows:

$$D_T \times D_H = ]T_{min}, T_{max}] \times ]0, 100], \quad T_{min} < T_{max} \tag{1}$$

The loss factor $C$ must satisfy:

C1) $C(t,h) \geq 1$ for every $(t,h) \in D_T \times D_H$.

C2) If $t \mapsto T_{min}$ then $C(t,h) \mapsto \infty$.

C3) If $t \mapsto T_{max}$ and $h \mapsto 0$ then $C(t,h) \mapsto 1$.

The first condition (C1) is determined by the model itself, while the remaining conditions are intrinsically linked to specific factors associated with the spread of wildfires, as depicted in Fig. 1. In essence, when the temperature is close to $T_{min}$, the expansion is unlikely to persist. However, if the land is suffering drought conditions ($h \approx 0$), the expansion is more prone to progress rapidly.

Humidity and temperature work in reverse: for small temperature values, the loss factor increases (as well as for large humidity values, although to a lesser extent). However, under conditions of extreme heat and dryness, the loss factor tends to be 1. We consider the function $f$ defined as follows:

$$f(t) = \frac{1}{(t - T_{min})^{\alpha}}, \quad \alpha \geq 0, \ t > T_{min} \tag{2}$$

Notice that if $t \mapsto T_{min}$ then $f(t) \mapsto \infty$. Moreover, $f(t) \geq 0, \forall t \in D_T$. Since the behaviour is reversed, we can consider $g$ as follows:

$$g(h) = h^{\beta}, \quad \beta \geq 0, \ h \in D_H \tag{3}$$

It can be observed that $g(h) \geq 0$ for every $h \in D_H$. So, we can define the loss factor function using functions $f$ and $g$ introduced previously.

$$C(t,h) := 1 + \gamma f(t)\, g(h) = 1 + \gamma \frac{h^{\beta}}{(t - T_{min})^{\alpha}}, \quad \alpha, \beta, \gamma \geq 0, \ (t,h) \in D_T \times D_H \tag{4}$$

**Property 2.1.** The loss factor $C(t,h)$ defined as in (4) satisfies:

1. *Conditions C1), C2), C3) are verified.*
2. *It is a function that depends on positive parameters $\alpha, \beta, \gamma$ such that*

$$\lim_{(\alpha,\beta) \mapsto (0,0)} C(t,h) = 1 + \gamma, \quad \forall (t,h) \in D_T \times D_H$$

Now, for the case of the infection probability $p_0$, our climate function must meet other characteristics. In this case:

P1) Since $p_0$ is a probability, then $p_0(t,h) \in [0,1], \ \forall (t,h) \in D_T \times D_H$.

P2) If $t \mapsto T_{min}$ then $p_0(t,h) \mapsto 0$.

P3) If $t \mapsto T_{max}$ and $h \mapsto 0$ then $p_0(t,h) \mapsto 1$.

In this sense, it is clear that the probability $p_0$ of infection increases as temperature increases and humidity decreases.

$$p_0(t,h) := \frac{1}{1 + \gamma f(t)\, g(h)} = \frac{1}{C(t,h)}, \quad \alpha, \beta, \gamma \geq 0, \ (t,h) \in D_T \times D_H \tag{5}$$

**Property 2.2.** The loss factor $p_0(t,h)$ defined as in (5) satisfies:

1. *Conditions P1), P2), P3) are verified.*
2. *It is a function that depends on positive parameters $\alpha, \beta, \gamma$ such that*

$$\lim_{(\alpha,\beta) \mapsto (0,0)} p_0(t,h) = \frac{1}{1 + \gamma}, \quad \forall (t,h) \in D_T \times D_H$$

#### 2.3.2. Increasing infection capacity

It should be noted that in (Boters-Pitarch et al., 2023a) we set the partitions to have 4 elements, which means that the scope was limited to 3 rows and/or columns. Due to the discontinuity in the expansion rates that can occur during a fire, limiting the range to only 3 rows or columns is a clear limitation of the model. So, it might be more appropriate to consider it as an adjustable element depending on the fire and its environmental conditions. Thus, we modify the Definition 3.4 in (Boters-Pitarch et al., 2023a) as follows:

**Definition 2.1.** Let $M \geq 2$ be an integer, and a partition $\mathscr{P} := \{u_i\}_{i=0}^{M+1}$ of the interval $[0,1]$ such that

$$0 = u_0 \leq u_1 \leq \ldots \leq u_M \leq u_{M+1} = 1$$

Then, if $i$ is the smallest positive integer satisfying

$$u_i \leq \rho \leq u_{i+1},$$

we say that the intensity affects up to the $i$-th layer.

**Remark 2.1.** Therefore, if $M = 3$, we have the same model. However, for particular cases, we could use $M > 3$ to get better estimations.

We can generalize Property 3.1 of (Boters-Pitarch et al., 2023a) by using the enlargement process explained in the same paper. Let $A$ be a matrix and let us define the norm $\|\cdot\|_{1,1}$ of $A$ as

$$\|A\|_{1,1} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |a_{ij}|$$

**Theorem 2.1.** *Suppose $(i,j) \in I_k$ is infected and $M \geq 1$. Let $\mathscr{P} := \{u_i\}_{i=0}^{M+1}$ be an arbitrary partition of the interval $[0,1]$, $p_0 \in [0,1]$ and*

$C \geq 1$. *Then, the marginal infectiousness E for our model is*

$$E(\mathscr{P}, C, p_0) = 8u_1 p_0 + \sum_{k=1}^{M} \frac{4}{\pi} (u_{k+1} - u_k) \int_0^{\frac{\pi}{4}} \|A_k\|_{1,1} \, d\theta \tag{6}$$

*Proof.* The proof is carried out reasoning as Property 3.1 in (Boters-Pitarch et al., 2023a). We only need to take into account that given an intensity $\rho$ for the wind, the expected infectiousness is the sum of the entries of the matrices $A_k$ of the enlargement process $\forall 1 \leq k \leq M+1$.□

Fig. 5 shows how $E$ increases as $M$ increases.

### 2.3.3. New update criterion

The Gumbel-Softmax function is a technique used in machine learning to approximate the selection of discrete elements in a differentiable manner. Based on the Gumbel distribution, which is a continuous distribution used for random sampling, the Gumbel-Softmax function applies the Gumbel transformation to a discrete distribution and then normalizes it using the Softmax function, resulting in a smooth and differentiable approximation of the discrete selection, see (Jang et al., 2016; Maddison et al., 2016).

The Gumbel-Softmax function is a more general function that can approximate any discrete distribution to a continuous distribution. Its usefulness in optimizing neural networks involving discrete variables, such as those used in reinforcement learning and generative models, can be seen in (Mena et al., 2018; Strypsteen and Bertrand, 2021). However, the Gumbel-Softmax function is not a perfect substitute for the binary Bernoulli random variable. Since the introduction of a stochastic element into the computational process allows us to keep that character, but it could have an impact on our estimates. Furthermore, the function introduces a temperature parameter $\tau$ that controls the level of approximation to the discrete variable. If $\tau$ tends to 0, the Gumbel-Softmax function tends to the discrete random variable, see Proposition 1 of (Maddison et al., 2016).

**Remark 2.2.** Mostly, the value of $\tau$ considered by default is $\tau = 1$.

Thus, if $m_{ij}^k$ refers to the state of the cell $(i,j)$ in the generation $k$, where

$$m_{ij}^k := \begin{cases} 0 & \text{if susceptible} \\ 1 & \text{if infected} \\ 2 & \text{if dead} \end{cases}$$

then, the new update criterion $\widetilde{\mathscr{U}}_k$ is defined by Eq. (7)

$$\begin{aligned} &\forall m_{ij}^k = 2 \Rightarrow m_{ij}^{k+1} = 2 \\ &\forall 1 \leq m_{ij}^k < 2 \Rightarrow m_{ij}^{k+1} = m_{ij}^k + \frac{1}{\Delta_{ID}} \\ &\forall m_{ij}^k = 0 \Rightarrow m_{ij}^{k+1} := \begin{cases} GS\left(p_{ij}^k, \tau\right) & \text{if} \quad 0 \leq p_{ij}^k < 1 \\ 1 & \text{if} \quad p_{ij}^k \geq 1 \end{cases} \end{aligned} \tag{7}$$

where $GS(\cdot, \tau)$ is the Gumbel-Softmax function and $\tau$ is the temperature parameter, which determines the approximation of this function to the real discrete distribution, see Proposition 1 of (Maddison et al., 2016).

## 3. Methodology

We apply our model to a real case, specifically the Beneixama fire, to

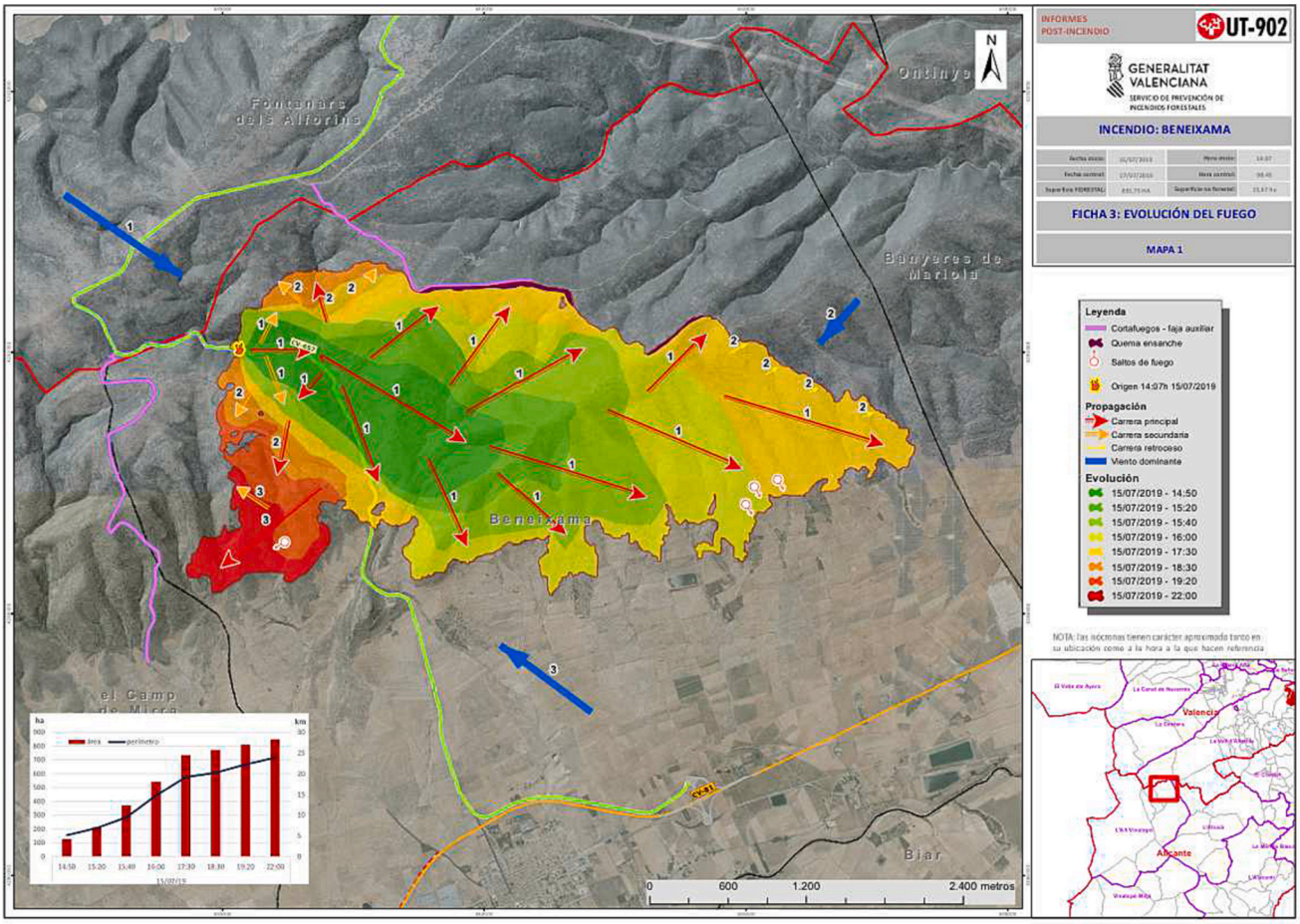



**Fig. 2.** A real forest fire caused in Beneixama, Alicante.

evaluate its ability to predict the evolution of the fire in a realistic scenario, see Fig. 2. Furthermore, we compare the results with the previous ones, before the modifications have been made.

The Beneixama wildfire has been selected as a case study due to some of its particular characteristics. First, it is a short-lived fire and this facilitates the analysis by avoiding the influence of prolonged environmental factors, such as the absence of sunshine, and minimizing the effects related to the chaotic behaviour of the model. Second, we have comprehensive monitoring of the fire, with a higher number of temporal measurements of the burned area compared to other similar fires. This provides us with a significant amount of information to test and evaluate the predictive capacity of our model in the evolution of the fire. Third, we have been able to obtain accurate weather data for this particular case, since they are provided by AVAMET (Valencian Association of Meteorology) and come from the station located in the area of the fire.

It is important to take into account the possible biases in the data used in our research. We need to distinguish between the data structures of the model. On the one hand, we have what we refer to as spatial inputs. For instance, the data extracted from Fig. 2 related to the states of each cell are a spatial input of dimensions $257 \times 257 \times 9$, denoted as $M$ (masks of states).

Since the fires are part of a larger database and have different scales, it has been necessary to unify the scales for a proper comparative analysis, see **Supplementary Material**. For that, data have been rescaled to a manageable size of $257 \times 257$ pixels, compatible with our computational platform, where each pixel is equivalent to one cell of the grid and every single cell refers to the same portion of land.

We have 9 two-dimensional arrays, which refer to the state $S_k$ of the fire with $k \in I$, where $I$ is the set of generations for which we have a real mask to compare the results. In case of having a two-dimensional representation of the slope or the existing vegetation type for each cell, these inputs would also be categorized as spatial inputs. Furthermore, it is important to mention that we can extract the initial state from $M$, considering it as the focal point of the fire. A time delay of 30 min has been defined as a generation period, which implies the use of a spline interpolation method to determine the climatic data according to actual values provided by the station. Then, since we have another data structure with wind, temperature, and humidity values stored for every generation, we can compute an estimation for each one.

**Remark 3.1.** In cases where data is available for the 30th minute or on the dot, the value provided by the station has been used directly, while in other cases an extrapolation has been made using the spline.

With our estimation and the representation $M$, we are able to assess and analyze the quality of our estimation for every $k \in I$.

To perform the analysis, we use the model as a Monte Carlo-based classification method, see Section 5 of (Boters-Pitarch et al., 2023b). This approach has the advantage of using various metrics, such as accuracy, F1-score, and confusion matrices, among others, to comprehensively assess the performance and effectiveness of our model.

Specifically, the use of confusion matrices will provide us with detailed information regarding both correct and incorrect classifications made by our model. This valuable insight will enable us to identify potential deficiencies in classifying specific classes or categories. In conclusion, we will present the evaluation results through the use of figures and numerical data. Graphical visualizations will effectively portray the performance estimation achievement.

## 4. Discussion and results

In this section, we present an exploratory analysis of the climatic data obtained during the development of the fire. Furthermore, we will assess the accuracy of the model and examine both its strengths and weaknesses in depicting the phenomenon. It is worth noting that the Beneixama fire is part of a broader database we are currently developing. To ensure comparability among the various registered fires, a transformation was necessary, so that each pixel from every fire refers to the same real-world terrain area. This involves a modification of the size of the fire mask. By adopting this methodology, the consistent applicability of the model to each fire in the database is ensured.

First of all, if we focus on the data related to the wind, we can highlight two fundamental aspects. In the early hours of the fire, the wind direction $\theta$ is from the southeast, belonging to the fourth quadrant. However, around 8:00 p.m. (12th generation), it undergoes a drastic change towards the northwest, belonging to the second quadrant. As for the intensity of the wind $\rho$, it remains very high throughout the fire's development, falling within the top decile of the intensity distribution in our geographical area, see Fig. 6. It is important to point out that we have calculated the distribution function of $\rho$ considering the following aspects:

i) We use climate data collected from the year 2018 to 2022, exclusively focusing on the summer months: June, July, and August.

ii) Data is collected from 7 meteorological stations located at various points throughout the Valencian territory, aiming to represent its diverse climate.

iii) Lastly, the analysis of its distribution allows us to determine that it follows an exponential distribution.

$$\rho \sim Exp(\lambda), \quad \lambda = 0.2175,$$

as shown in Fig. 7.

In Fig. 6, we observe the temperature and humidity conditions favour the spread of the fire in its first hours, with temperatures above 34 degrees and relative humidity below 25%. However, after 8 p.m. these conditions are less favorable due to both a reduction of temperatures and an increase of relative humidity.

In Fig. 8, we can see how the original shape of the mask representing

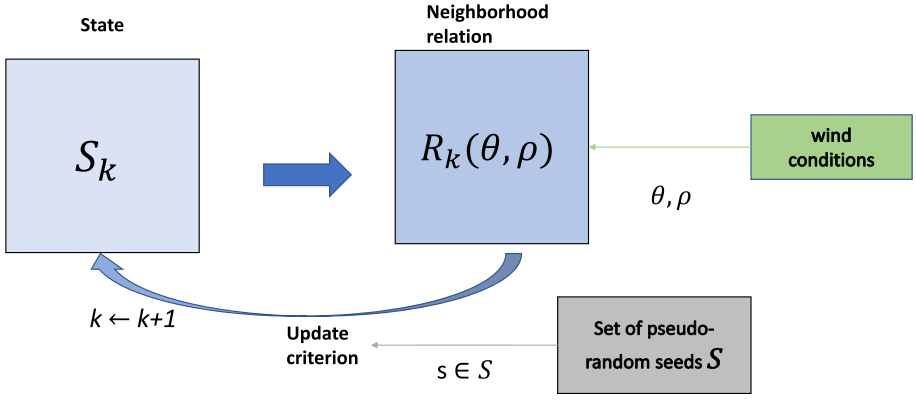



**Fig. 3.** Overview of the operational principles of the proposed model.

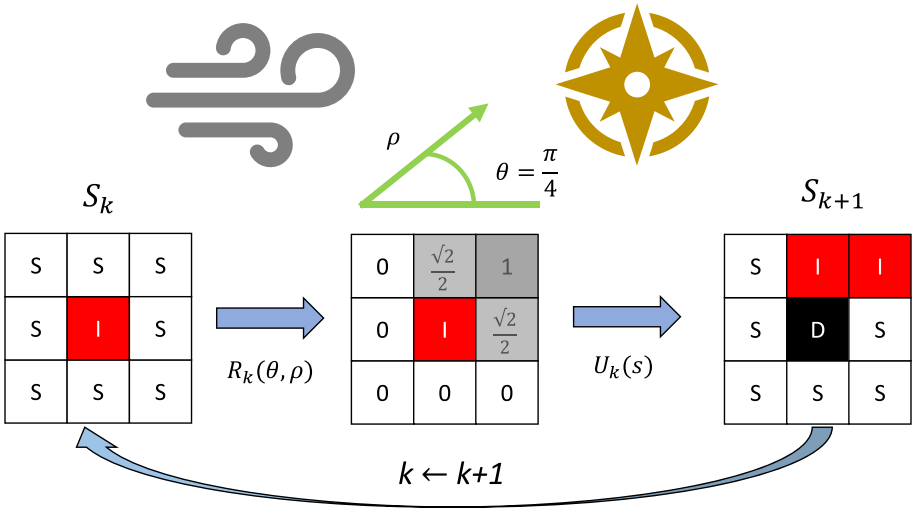



**Fig. 4.** Performance framework of the proposed model.

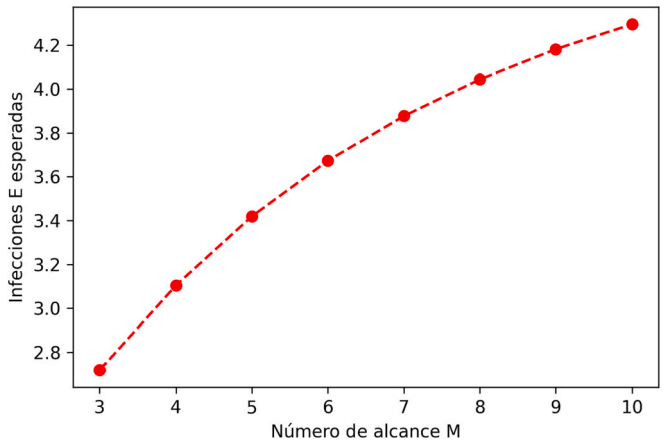



**Fig. 5.** Infectious capacity $E$ as a function of $M$ for partitions with equidistant elements.

the fire is maintained, as shown in Fig. 2. Hence, it is essential to note that the unburnt elements represent 97% of the sample, while the burnt or affected elements only represent 3%.

**Remark 4.1**. From now on, healthy, i.e. unburned cells will be considered as positive. On the other hand, burned (dead) or affected (infected) cells will be considered negative.

This imbalance poses a problem when interpreting the accuracy statistics of the model, as it can distort the values because of over-represented class 0, see (He and Garcia, 2009). Thus, in Section 4.1, we will go into more detail on this subject.

It also shows the evolution of the burned cells during the forest fire. It can be seen that the burned area grows fast at the beginning, but as the fire spreads, its growth rate decreases. Possibly human action in extinguishing the fire, as well as the presence of geographical obstacles or firebreaks designed to stop the advance of the fire, may have influenced this pattern. However, it is important to mention that our model has not taken these variables into account, and this will require a critical analysis of the model results in its final stages.

### 4.1. Analysis of model results

As previously mentioned, due to the unbalance between classes in our data, it is necessary to use alternative measures to assess the accuracy of our model. In this way, the performance of the model is not distorted by this aspect.

We consider the false positive (**FP**) and false negative (**FN**) percentage to measure how good the model fit is. These metrics allow us to know the percentage of hits that our model achieves for each class. In addition, we have observed that the metrics such as balanced accuracy (**B. Acc**), F-measure $F_\beta$, Cohen's kappa coefficient $k$ and Matthew's correlation coefficient (**MCC**) have proven to be adequate in similar contexts, although they have their own limitations (Chicco et al., 2021; Chicco and Jurman, 2020; Lee et al., 2021). For this reason, Table 1 presents the results of these metrics in relation to the parameters of our model and the final results obtained in our real fire. Using these alternative measures gives us a more complete and accurate picture of the performance of our model, which is essential to properly evaluate.

We use the following partitions, which determine the importance of $\rho$ in our estimations:

$$\mathscr{P}_0 := \{0, 0.1, 0.5, 0.9, 1\}$$
$$\mathscr{P}_1 := \{0, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$$
$$\mathscr{P}_2 := \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$$

It can be seen from Table 1 that the model results for partition $\mathscr{P}_0$ lack range capability. This is justified by **FN**, since for the smallest value of $C$ (which provides the maximum range for that partition), the best result is obtained in this metric, with an error of about 40%. In other words, we fail to accurately account for the burned cells by means of these range parameters.

On the contrary, this is not the case for partitions $\mathscr{P}_1$ and $\mathscr{P}_2$, thanks to the increased influence of wind intensity by increasing the number of exposed cells per infected cell, see Section 2.3.2. In these cases, it can be observed that the best result in terms of **FN** is achieved for values of C greater than 1, indicating that a decrease in the extent of these partitions also decreases the amount of **FN**. Moreover, in both cases, a **B. Acc** above 91%, which together with $\kappa$ and **MCC** leads us to affirm the existence of a strong dependence relationship between the results of our model and that of the fire.

Fig. 9 shows the comparison between the fire mask and the best results of our model for $\mathscr{P}_1$ and $\mathscr{P}_2$, with values of $C$ equal to 1.5 and 2, respectively.

Thus, we consider the estimate provided by partition $\mathscr{P}_1$ to be of higher quality than that provided by $\mathscr{P}_2$, as we can see in Fig. 9. Moreover, although the **B. Acc** is better for $\mathscr{P}_2$, worse results are obtained for the $F_\beta$, $k$ and **MCC** metrics. This fact is due to the overreaching of the model for that partition, which leads to good results in terms of **FN**, at the expense of a worsening of **FP**, and given the predominance of

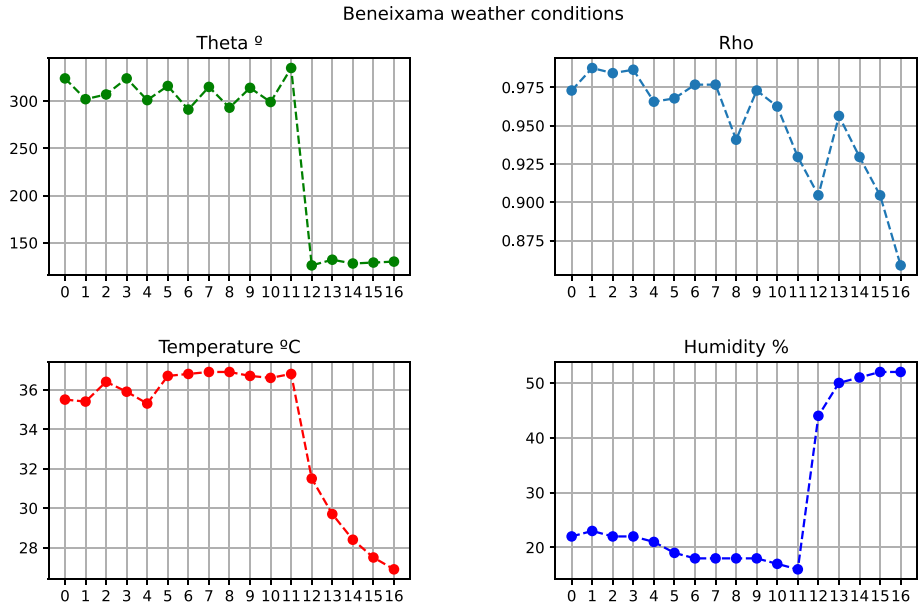



**Fig. 6.** Meteorological conditions during the forest fire: Wind direction (upper-left), Wind power (upper-right), Temperature (lower-left), Relative humidity (lower-right).

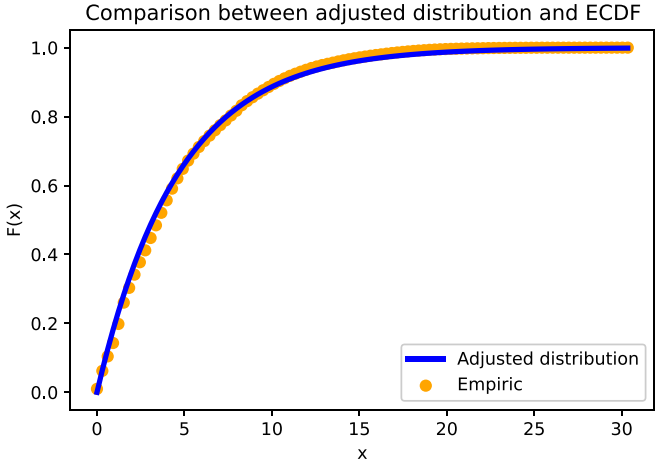



**Fig. 7.** Comparison between the estimated and the empirical distribution function.

the positive class, we obtain such a worsening in the evaluation metrics and correlation coefficients. Then, from now on, we will turn our attention to the results of our model for the $\mathscr{P}_1$ partition.

At the same time, it is highlighted that our model has the capability to function as a Monte Carlo-based classification method, as detailed in (Boters-Pitarch et al., 2023b). In Fig. 10, both the heat map representing the active focus in the final state and the graphical representations of the areas of land that have been affected by the fire are presented. However, in the case of the heat map illustrating the areas of burnt land, certain flaws inherent in the construction of our model are evident.

These flaws are attributable to the way in which the neighbourhood relationship is implemented by quadrants in our approach. In other words, a fire-infected cell can spread fire to healthy cells belonging to one of the quadrants. This dynamic can result in estimates with very sharp corners, especially when the direction $\theta$ of the wind remains in the same quadrant constantly. At the same time, as mentioned above, our model is not able to take into account the firebreak areas, which allows the expansion over the north-east of the forest when in reality, this would not be possible, see Fig. 2. In other words, our model overestimates the potential burnt area in this case.

Our model has a key advantage in leveraging previous states to

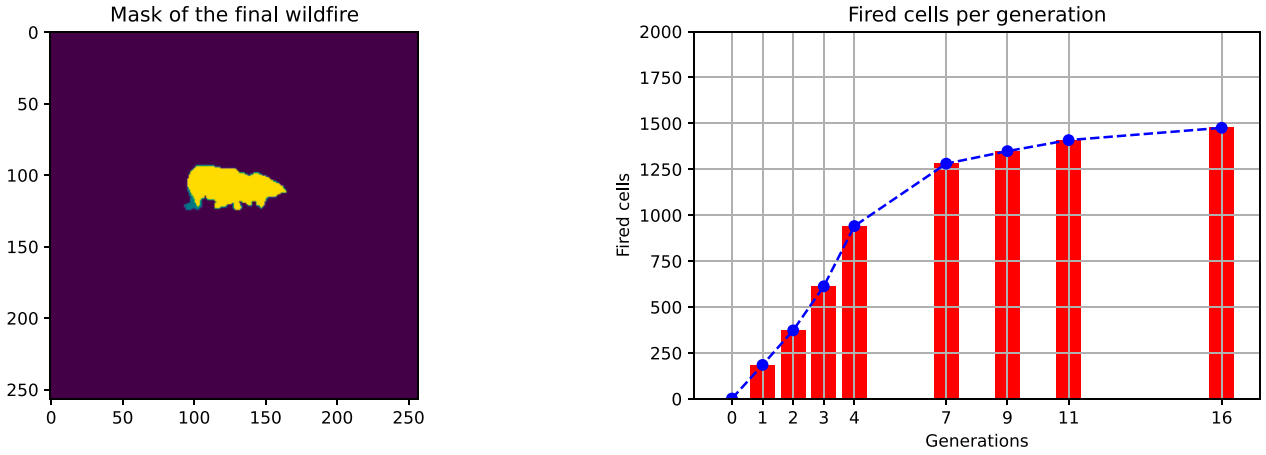



**Fig. 8.** The final mask of the model (left). Number of fired cells per generation (right).

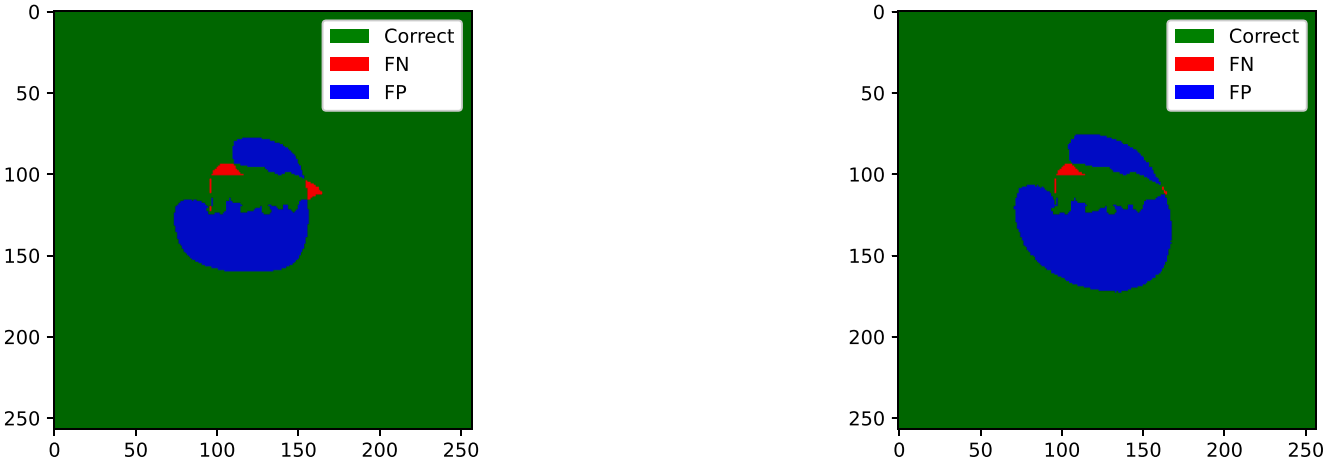



**Fig. 9.** Comparison between the fire mask and the best model results for partitions $\mathscr{P}_1$ (left) and $\mathscr{P}_2$ (right).

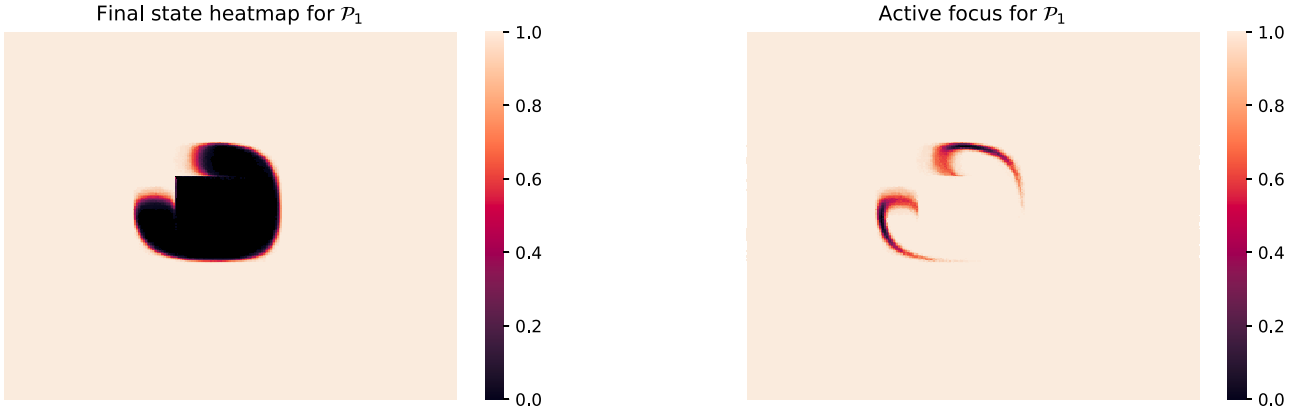



**Fig. 10.** Heat map of the final state of the fire (left). Active hot spots in the final state of the fire (right).

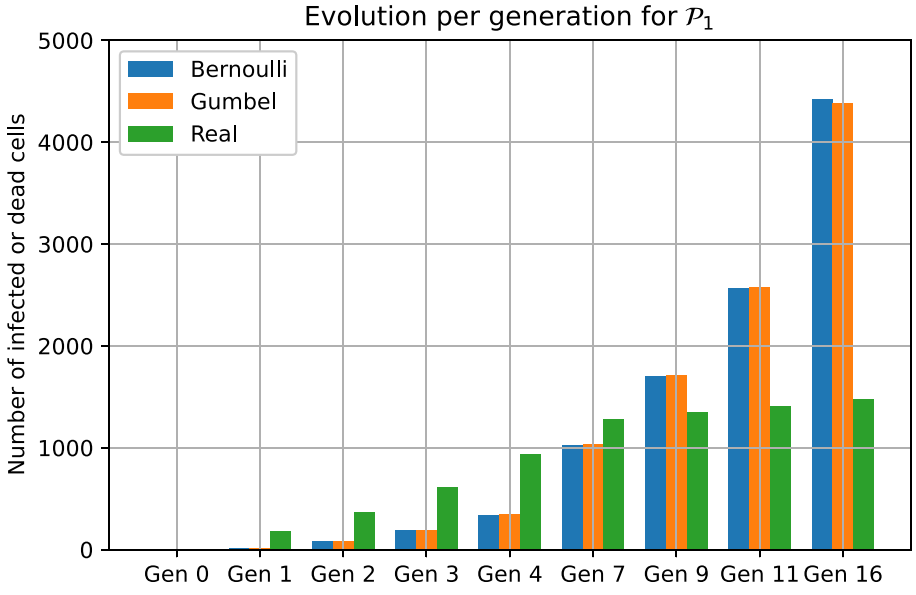



**Fig. 11.** Temporal evolution of the number of burnt cells calculated by the model.

calculate subsequent states, which provides a valuable approach to tracking and studying outcomes over time. Despite this strength, our model faces challenges in accurately representing fire initiation, see Fig. 11. This initial difficulty has a noticeable effect on all subsequent estimates given the recursive nature of the model. The ability to capture the initial rate and speed of the fire is critical to accurately forecast its trajectory and magnitude.

This limitation leads us to adapt our neighbourhood relationship. A

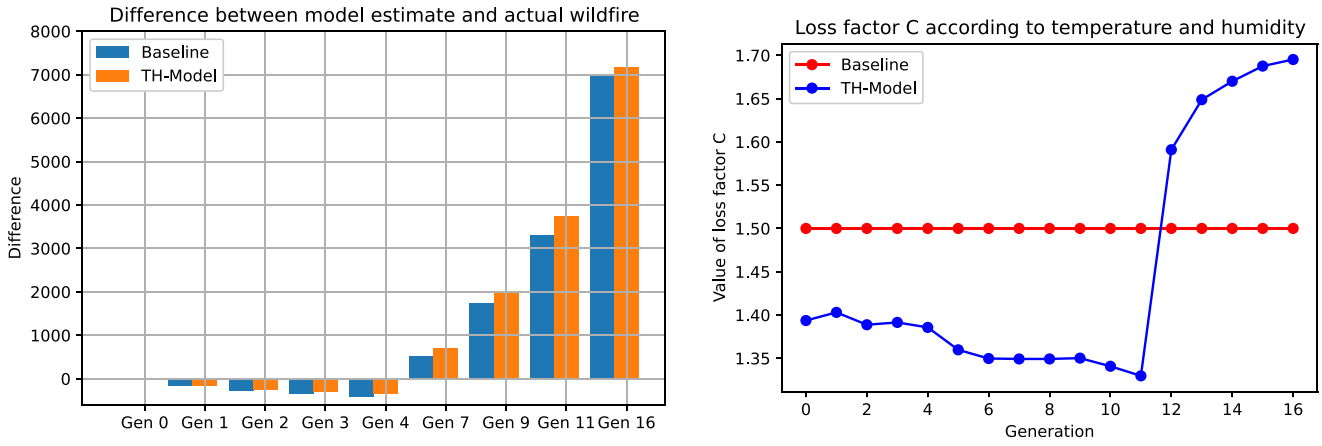



**Fig. 12.** Differences between our estimations and real fire (left). Values taken by the loss factor $C$ as a function of the fire climate (right).

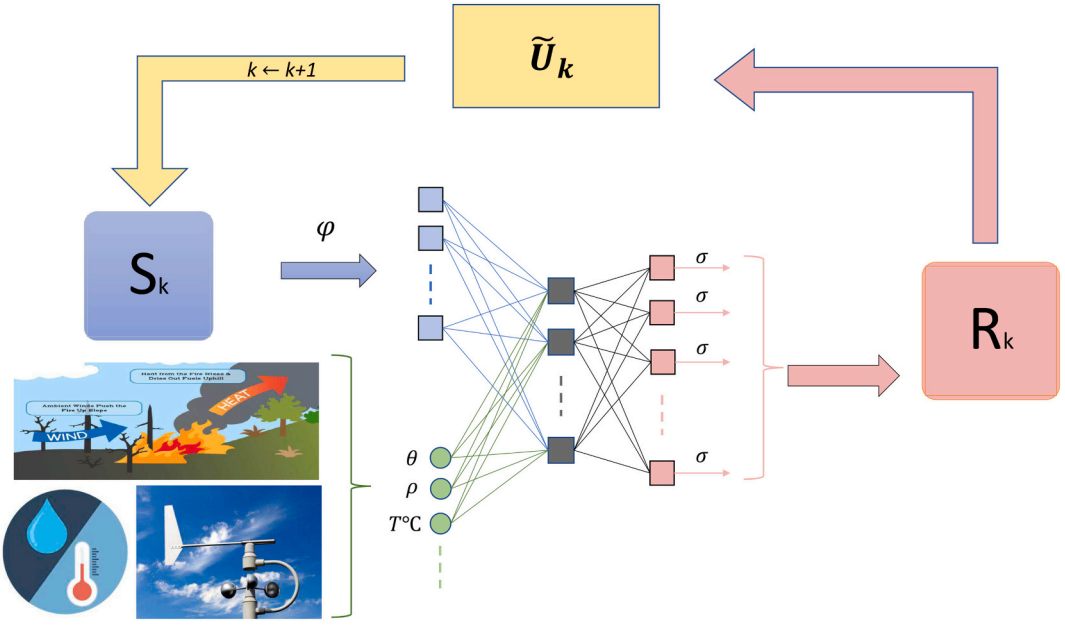



**Fig. 13.** New architecture to learn neighbourhood relations.

more flexible and discontinuous neighbourhood relationship could be the key to capture the various trends that emerge during fires. In addition, introducing other variables such as the influence of human actions or the existence of firebreaks could enrich our model and improve its predictive capacity, in the final states, by reflecting real-world factors that influence fire development.

It is important to note that, although our neighbourhood relationship $R^k$ is adjustable, its scope is limited and depends exclusively on $\theta$ and $\rho$. This limitation leads us to recognise the presence of other significant variables that shape the extent of the fire. Beyond $\theta$ and $\rho$, exploring and considering other dimensions in our modelling becomes a crucial step towards a more complete and accurate understanding of wildfires, see Subsection 2.3.1 and Section 5.

### 4.2. Analysis of model using temperature and humidity

The main objective of the following section is to assess the impact caused by the enhancements detailed in Section 2.3.1. In these enhancements, parameters undergo a conversion into climatic functions, with variables based on temperature and humidity.

The outcomes of these adjustments are showcased in Table 2, they represent the results of the aforementioned metrics for the modified model (TH-model). The $\mathscr{P}_1$ partition is used as a reference point for the intensity of the wind. A thorough analysis of these outcomes reveals an average increment of 4 points **B.Acc** metric. Generally, this entails a heightened convergence in results and a reduced sensitivity to fluctuations in the loss factor $C$.

However, an important inquiry arises as to whether these enhancements also contribute to a more precise representation of fire evolution. By closely examining Fig. 12, the variability in the $C$ values across generations becomes apparent. These values exhibit a tendency to be lower in the initial stages of the process, resulting in a broader reach of the model during its initial states and higher values in the final stages. This modification effectively addresses previously discussed limitations concerning the rate of fire expansion. This effect is a direct outcome of the influence of climatic data (temperature and humidity) on the increase of the loss factor.

Consequently, the implemented improvements, while having generated a moderate impact due to the implications of parameters in fire expansion (see Section 3 of (Boters-Pitarch et al., 2023b)), have consistently and successfully followed the logic and objectives outlined in their implementation, as we can observe seeing generation 11 and 16,

**Table 1**
Results of model evaluation metrics for different partitions and $C$.

| Parameters | | Metrics | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Partitions** | **Loss Factor** | **FP** | **FN** | **B. Acc** | $F_{0.5}{}^{*}$ | $F_1{}^{*}$ | $F_2{}^{*}$ | $k^{*}$ | **MCC*** |
| $\mathscr{P}_0$ | $C = 1$ | 1.59 | 39.23 | 79.59 | 48.88 | 52.75 | 57.29 | 51.52 | 52.00 |
| | $C = 1.5$ | 1.35 | 39.43 | 79.61 | 52.32 | 55.13 | 58.27 | 54.01 | 54.24 |
| | $C = 2$ | 0.92 | 45.19 | 76.95 | 57.07 | 56.20 | 55.36 | 55.22 | 55.24 |
| | $C = 2.5$ | 0.63 | 51.22 | 74.08 | 60.24 | 55.36 | 51.22 | 54.48 | 55.01 |
| | $C = 3$ | 0.44 | 56.10 | 71.73 | 62.16 | 53.78 | 47.38 | 52.96 | 54.41 |
| | **Average** | 0.99 | 46.23 | 76.39 | 56.13 | 54.64 | 53.90 | 53.64 | 54.18 |
| $\mathscr{P}_1$ | $C = 1$ | 6.23 | 14.97 | 89.40 | 27.79 | 37.17 | 56.12 | 34.89 | 42.96 |
| | $C = 1.5$ | 5.54 | 11.99 | **91.23** | 30.95 | 40.89 | 60.24 | 38.79 | 46.61 |
| | $C = 2$ | 3.40 | 21.61 | 87.49 | 38.84 | 47.90 | 62.48 | 46.23 | 50.47 |
| | $C = 2.5$ | 1.97 | 33.13 | 82.45 | 46.93 | 52.84 | 60.45 | 51.53 | 52.77 |
| | $C = 3$ | 1.19 | 42.34 | 78.23 | 53.56 | 55.03 | 56.57 | 53.95 | 54.01 |
| | **Average** | 3.67 | 24.81 | 85.76 | 39.61 | 46.77 | 59.17 | 45.08 | 49.36 |
| $\mathscr{P}_2$ | $C = 1$ | 24.28 | 16.12 | 79.80 | 8.95 | 13.46 | 27.12 | 9.75 | 20.18 |
| | $C = 1.5$ | 19.39 | 9.35 | 85.63 | 11.76 | 17.45 | 33.85 | 13.98 | 25.87 |
| | $C = 2$ | 8.06 | 6.30 | **92.82** | 24.85 | 34.30 | 55.35 | 31.81 | 42.24 |
| | $C = 2.5$ | 3.37 | 22.56 | 87.04 | 38.75 | 47.68 | 61.97 | 46.01 | 50.12 |
| | $C = 3$ | 1.66 | 36.65 | 80.84 | 49.15 | 53.66 | 59.08 | 52.43 | 53.09 |
| | **Average** | 11.35 | 18.20 | 85.23 | 26.69 | 33.31 | 47.47 | 30.80 | 38.30 |

Coefficients with * are multiplied by 100.

**Table 2**
Results of evaluation metrics for TH-model, $\mathscr{P}_1$ and parameters based on temperature and humidity.

| $\alpha$ | $\beta$ | $\gamma$ | **FP** | **FN** | **B. Acc** | $F_{0.5}{}^{*}$ | $F_1{}^{*}$ | $F_2{}^{*}$ | $k^{*}$ | **MCC*** | $\overline{C}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.5 | 0.5 | 0.5 | 5.62 | 13.41 | 90.48 | 30.26 | 40.03 | 59.09 | 37.89 | 45.62 | 1.45 |
| | | 1.0 | 3.82 | 18.70 | 88.74 | 37.15 | 46.65 | 62.68 | 44.89 | 50.00 | 1.90 |
| | | 1.5 | 2.48 | 28.18 | 84.67 | 43.69 | 51.21 | 61.86 | 49.76 | 52.09 | 2.36 |
| 1.0 | 0.5 | 0.5 | 6.32 | 14.30 | 89.69 | 27.66 | 37.08 | 56.22 | 34.80 | 43.02 | 1.08 |
| | | 1.0 | 6.29 | 14.02 | 89.84 | 27.83 | 37.29 | 56.48 | 35.01 | 43.24 | 1.16 |
| | | 1.5 | 6.20 | 13.96 | 89.92 | 28.13 | 37.62 | 56.80 | 35.37 | 43.53 | 1.24 |
| | 1.0 | 0.5 | 5.49 | 15.24 | 89.63 | 30.27 | 39.89 | 58.45 | 37.76 | 45.14 | 1.45 |
| | | 1.0 | 4.10 | 17.95 | 88.97 | 35.81 | 45.41 | 62.03 | 43.58 | 49.13 | 1.89 |
| | | 1.5 | 3.02 | 23.78 | 86.60 | 40.86 | 49.46 | 62.66 | 47.89 | 51.37 | 2.34 |
| 1.5 | 0.5 | 0.5 | 6.27 | 14.70 | 89.51 | 27.71 | 37.11 | 56.14 | 34.83 | 42.96 | 1.01 |
| | | 1.0 | 6.29 | 14.63 | 89.54 | 27.68 | 37.08 | 56.12 | 34.79 | 42.95 | 1.03 |
| | | 1.5 | 6.29 | 14.57 | 89.57 | 27.70 | 37.10 | 56.17 | 34.82 | 42.98 | 1.04 |
| | 1.0 | 0.5 | 6.27 | 14.50 | 89.62 | 27.78 | 37.19 | 56.27 | 34.92 | 43.07 | 1.08 |
| | | 1.0 | 6.17 | 14.91 | 89.46 | 27.98 | 37.39 | 56.34 | 35.13 | 43.16 | 1.16 |
| | | 1.5 | 6.03 | 14.84 | 89.57 | 28.47 | 37.95 | 56.86 | 35.71 | 43.63 | 1.24 |
| | 1.5 | 0.5 | 5.25 | 16.19 | 89.28 | 30.93 | 40.52 | 58.71 | 38.43 | 45.48 | 1.48 |
| | | 1.0 | 4.24 | 17.48 | 89.14 | 35.20 | 44.85 | 61.76 | 42.99 | 48.76 | 1.96 |
| | | 1.5 | 3.37 | 20.87 | 87.88 | 39.29 | 48.43 | 63.13 | 46.78 | 51.05 | 2.44 |
| | **Average** | | 5.20 | 16.79 | 89.01 | 31.91 | 41.24 | 58.76 | 39.19 | 45.95 | 1.52 |

Coefficients with * are multiplied by 100.

and how we have improved the rate of expansion in the initial stages and how it has been greatly moderated in the final stages.

Despite the positive outcomes yielded by our model, we contend that substantial improvements could be achieved by implementing a more flexible and complete neighbourhood relation $R_k$ in terms of variables. For this reason, in Section 5, we explore an architecture capable of learning such neighbourhood relationships through experience and accumulated data.

## 5. Intelligent cellular scheme

In this section, we aim to present an innovative architecture that uses existing data and historical databases to generalize the concept of neighbourhood relations. To achieve it, machine learning is a powerful tool, with a backpropagation algorithm (Rumelhart et al., 1986) at the heart of it.

However, for machine learning algorithms like backpropagation to be successful, a **critical condition** must be met: all functions involved **must be differentiable** to compute gradients accurately. As said above, to make the update criterion differentiable, we use the Gumbel softmax function. So, by changing our update criterion to $\widetilde{U}_k$, we manage to compute the gradients.

This new data-driven architecture aims to learn the neighbourhood relationship $R_k$ to improve the results of our model. In addition, learning the neighbourhood relationship has value in itself, as it allows us to know and learn how the different elements of the system interact. For this purpose, as we can see in Fig. 13 a neural network **ANN** (or other objects) might be defined as $R_k$, which may take two kinds of inputs:

- **Spatial inputs**: The previous architecture could use variables such as the previous state $S_k$, the slope or even the plant fuel (composition of the vegetation that serves as the primary source of material that can burn and sustain a wildfire) for each cell.
- **Numerical inputs**: At the same time, it also could take into account as many numerical variables as we want such as temperature, humidity and intensity of the wind among others.

This proposal presents several significant advantages. Firstly, it facilitates the inclusion of new climate variables in the learning and computation of the neighbourhood relationship as we only need to add them as inputs to the **ANN**. Secondly, we do not impose any limits (in the scope) or fix any characteristics or appearances (corners) of the infection

function. We also avoid defining parameters, this allows us to learn non-linear relationships with an infection capacity greater than those established in (Boters-Pitarch et al., 2023a). Thirdly, the $\varphi$ filter ensures that the expansion is carried out only by the infected elements of the system, making the $\Delta_{ID}$ parameter decisive during learning, which represents a significant advance over the previous relationship. Finally, using the sigmoid activation function $\sigma$ (or similar) as the output of the network guarantees that all the cells of the neighbourhood relationship can be interpreted as probabilities, without having to consider those with a probability greater than one, i.e. $R_k \in [0,1]^{N\times,N}$.

Thus, setting a random seed $s$, we can compute the estimation $\widehat{y}^s_{k+1}$ of the model in the generation $k+1$ by reasoning as Fig. 13

$$\widehat{y}^s_{k+1} \leftarrow S_0 R_0 \widetilde{U}_0 \ldots S_k R_k \widetilde{U}_k \tag{8}$$

and then an estimation $\widehat{y}_{k+1}$, based on the Monte Carlo-based classification method of (Boters-Pitarch et al., 2023b), can be calculated by averaging the outcomes of different random seeds.

Hence, let us consider a loss function $F$, which aims to evaluate the error between our estimate mask $\widehat{y}_i$ and the real mask $y_i$ for all generation $i \in I$, where $I$ is the set of generations for which we have a real mask to compare the results. In this sense, we could define a discount parameter $\widetilde{\gamma} \in [0,1]$ to guide the learning process, since by adjusting this parameter, it is possible to control and seek a balance between the relative importance of adjusting better in the early and/or late stages. Thus, the loss result $L_{\widetilde{\gamma}}$ is computed as:

$$L_{\widetilde{\gamma}} = \sum_{\substack{i\in I \\ j=0}}^{|I|} \widetilde{\gamma}^j F(y_i, \widehat{y}_i) \tag{9}$$

Therefore, to determine the optimal neighbourhood relation $R^k$, we must find the architecture parameters $\alpha_1, \ldots, \alpha_M$ that minimize the value of $L_{\widetilde{\gamma}}$, i.e.

$$\alpha_1, \ldots, \alpha_M : \ arg\,min L_{\widetilde{\gamma}}$$

which implies that our estimates $\widehat{y}_i$ optimally fits the real masks $y_i$ taking into account the discount factor $\widetilde{\gamma}$ and for all generations $i$ in set $I$.

## 6. Conclusions

In this study, we have performed a comprehensive evaluation of our previously developed model (Boters-Pitarch et al., 2023a; Boters-Pitarch et al., 2023b). Some limitations were detected which could jeopardize future improvements. Among them, we can highlight that the scope of the infection is bounded by the parameters and consequently, the quality of our results depends on the suitable selection of them. At the same time, the model shows a marked sensitivity to wind patterns as a climate variable. This, together with our method of updating the model by random draws, introduces complexity to the model fitting process. For each of the above limitations, we have proposed solutions such as: modifying the size of the partition parameter to increase the infection capacity, allowing the use of other variables, and changing the update function of the grid states. After that, we have also compared the outcomes of the initial model with those of the improved version in a realistic scenario, the Beneixama fire incident.

Despite the improvements and favorable results, we acknowledge the complexity of refining the model and reaching more accurate results within the current theoretical framework. This is because the existing approach may not be adequately versatile or comprehensive in terms of variables to achieve highly accurate fire tracking. Consequently, in response to this challenge, we have introduced a novel architecture in Section 5 that integrates the application of machine learning into our model. This architecture keeps its nature while introducing an intelligent element to learn the interaction between the cells in the system.

The main contribution of this new strategy is to enable the use of climatic data and spatial distribution of the fire to calculate the neighbouring relationships. So, this architecture allows learning and therefore understanding how the cells interact with each other. In other words, it aims to determine with greater precision the cells potentially affected in each iteration, from the climatic conditions and their prior state. We firmly believe that this data-driven approach can provide a powerful tool for enhancing the predictive capability of our model and moving towards a more accurate and effective monitoring of forest fires. Furthermore, given the generic approach of our architecture, we believe that it can adapt to other phenomena that combine spatial and climatic data, and whose final result is strictly dependent on the previous ones.

Finally, future work will aim to collect enough climatic and spatial data to effectively implement the proposed new strategy.

## CRediT authorship contribution statement

**Joan Boters Pitarch:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **María Teresa Signes Pont:** Validation, Supervision, Resources, Project administration. **Julian Szymański:** Validation, Supervision. **Higinio Mora Mora:** Validation, Supervision.

## 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.

## Acknowledgement

This research is funded by Generalitat Valenciana, project AICO/2021/331. This work does not have any conflicts of interest.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2023.102456.

## References

Alexandridis, A., Vakalis, D., Siettos, C.I., Bafas, G.V., 2008. A cellular automata model for forest fire spread prediction: the case of the wildfire that swept through spetses island in 1990. Appl. Math. Comput. 204 (1), 191–201.
Amat, Sergio, Aràndiga, Francesc, Arnau, José Vicente, Donat Beneito, Rosa M., Mestre, Pep Mulet, Sancho, Rosa Peris, 2002. Aproximació numèrica, vol. 60. Universitat de València.
Boccaletti, Stefanos, Celso Grebogi, Y.-C., Lai, Hector Mancini, Maza, Diego, 2000. The control of chaos: theory and applications. Phys. Rep. 329 (3), 103–197.
Boters-Pitarch, Joan, Signes-Pont, María Teresa, Szymanski, Julian, Mora-Mora, Higinio, 2023a. A new stochastic approach to the spread of environmental events enhanced by the wind. Math. Methods Appl. Sci. 1-9, 1–9.
Boters-Pitarch, Joan, Signes-Pont, María Teresa, Szymanski, Julian, Mora-Mora, Higinio, 2023b. Application of a stochastic compatmental model to approach the spread of environmental events with climatic bias. Eco. Inform. 77, 102266.
Chicco, Davide, Jurman, Giuseppe, 2020. The advantages of the Matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. BMC Genomics 21 (1), 1–13.
Chicco, Davide, Warrens, Matthijs J., Jurman, Giuseppe, 2021. The Matthews correlation coefficient (mcc) is more informative than cohen's kappa and brier score in binary classification assessment. IEEE Access 9, 78368–78381.
Cruz, Miguel G., Alexander, Martin E., Sullivan, Andrew L., Gould, James S., Kilinc, Musa, 2018. Assessing improvements in models used to operationally predict wildland fire rate of spread. Environ. Model. Softw. 105, 54–63.
Edwards, Paul N., 2011. History of climate modeling. Wiley Interdiscip. Rev. Clim. Chang. 2 (1), 128–139.

Emmanuel, Tlamelo, Maupong, Thabiso, Mpoeleng, Dimane, Semong, Thabo, Mphago, Banyatsang, Tabona, Oteng, 2021. A survey on missing data in machine learning. J. Big Data 8 (1), 1–37.
Földi, László, Kuti, Rajmund, 2016. Characteristics of forest fires and their impact on the environment. Acad. Appl. Res. Military Public Manag. Sci. 15 (1), 5–17.
Ganteaume, Anne, Camia, Andrea, Jappiot, Marielle, San-Miguel-Ayanz, Jesus, Long-Fournel, Marlène, Lampin, Corinne, 2013. A review of the main driving factors of forest fire ignition over europe. Environ. Manag. 51, 651–662.
He, Haibo, Garcia, Edwardo A., 2009. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 21 (9), 1263–1284.
Hernández Encinas, A., Hernández Encinas, L., Hoya White, S., Martín, A., del Rey, and G Rodríguez Sánchez., 2007a. Simulation of forest fire fronts using cellular automata. Adv. Eng. Softw. 38 (6), 372–378.
Hernández Encinas, L., Hoya White, S., Martín Del Rey, A., Rodríguez, G., Sánchez., 2007b. Modelling forest fire spread using hexagonal cellular automata. Appl. Math. Model. 31 (6), 1213–1227.
Jang, Eric, Gu, Shixiang, Poole, Ben, 2016. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*.
Karafyllidis, Ioannis, Thanailakis, Adonios, 1997. A model for predicting forest fire spreading using cellular automata. Ecol. Model. 99 (1), 87–97.
Kermark, M., Mckendrick, A., 1927. Contributions to the mathematical theory of epidemics. Part I. Proc. R. Soc. A 115 (5), 700–721.
Kuang, Yang, Nagy, John D., Eikenberry, Steffen E., 2018. Introduction to Mathematical Oncology. CRC Press.
Lee, Namgil, Yang, Heejung, Yoo, Hojin, 2021. A surrogate loss function for optimization of $F_\beta$ score in binary classification with imbalanced data. *arXiv preprint arXiv: 2104.01459*.
Maddison, Chris J., Mnih, Andriy, Teh, Yee Whye, 2016. The concrete distribution: A continuous relaxation of discrete random variables. *arXiv preprint arXiv:1611.00712*.
Malthus, Thomas Robert, Winch, Donald, James, Patricia, 1992. Malthus:'An Essay on the Principle of Population'. Cambridge University Press.
Mena, Gonzalo, Belanger, David, Linderman, Scott, Snoek, Jasper, 2018. Learning latent permutations with gumbel-sinkhorn networks. *arXiv preprint arXiv:1802.08665*.
Ntinas, Vasileios G., Moutafis, Byron E., Trunfio, Giuseppe A., Georgios, Ch., Sirakoulis., 2017. Parallel fuzzy cellular automata for data-driven simulation of wildfire spreading. J. Comput. Sci. 21, 469–485.
Rumelhart, David E., Hinton, Geoffrey E., Williams, Ronald J., 1986. Learning representations by back-propagating errors. Nature 323 (6088), 533–536.
Saar-Tsechansky, Maytal, Provost, Foster, 2007. Handling Missing Values when Applying Classification Models.
Sistema integrado de gestión de incendios forestales: Informes post-incendio, 2023. `prevencionincendiosgva.es/Incendios/IncendiosInformesPostIncendioList`. Accessed on 24 March.
Stojanova, Daniela, Panov, Panče, Kobler, Andrej, Džeroski, Sašo, Taškova, Katerina, 2006. Learning to predict forest fires with different data mining techniques. In: Conference on Data Mining and Data Warehouses (SiKDD 2006), Ljubljana, Slovenia, pp. 255–258.
Strypsteen, Thomas, Bertrand, Alexander, 2021. End-to-end learnable eeg channel selection for deep neural networks with gumbel-softmax. J. Neural Eng. 18 (4), 0460a9.
Wangersky, Peter J., 1978. Lotka-volterra population models. Annu. Rev. Ecol. Syst. 9 (1), 189–218.
Yang, Xubing, Hua, Zhichun, Zhang, Li, Fan, Xijian, Zhang, Fuquan, Ye, Qiaolin, Liyong, Fu., 2023. Preferred vector machine for forest fire detection. Pattern Recogn. 109722.
Ying, Lingxiao, Han, Jie, Yongsheng, Du, Shen, Zehao, 2018. Forest fire characteristics in China: spatial patterns and determinants with thresholds. For. Ecol. Manag. 424, 345–354.