

# Modeling the spatial evolution wildfires using random spread process

Carlos Díaz-Avalos<sup>1</sup> | Pablo Juan<sup>2,3</sup> 

<sup>1</sup>UNAM, Mexico

<sup>2</sup>IMAC, Instituto Universitario de Matemáticas y Aplicaciones de Castellón, Spain

<sup>3</sup>GRECS, Universitat de Girona, C/ de la Universitat de Girona 10, Girona, Spain

## Correspondence

Pablo Juan, IMAC, Instituto Universitario de Matemáticas y Aplicaciones de Castellón, Spain.

Email: [juan@uji.es](mailto:juan@uji.es)

## Funding information

PAPIIT-DGAPA, Grant/Award Number: IG100221

## Abstract

The study of wildfire spread and the growth of the area burned is an important task in ecological studies and in other contexts. In this work we present a model for fire spread and show the results obtained from simulations of burned areas. The model is based on probabilities of fire at different locations. Such probabilities are obtained from the intensity function of a spatial point process model fitted to the observed pattern of fires in the Valencian Community for the years 1993–2015. The models, applied to different wildfires in Spain, including the different temporal states, combines the features of a network model with those of a quasi-physical model of the interaction between burning and nonburning cells, which strongly depends on covariates. The results of the simulated wildfire burned areas resemble the burned areas observed in real cases, suggesting that the model proposed, based on a Markov process called Random Spread Process, works adequately. The model can be extended to simulate other random spread processes such as epidemics.

## KEYWORDS

burned area, random spread process, valencian communisty, wildfires

## 1 | INTRODUCTION

Landscape patterns are determined by the frequency, intensity and extent of disturbances. In this regard, wildfires play an important role because they have the potential to affect severely the forest dynamics and to produce geo-hydrological changes. At landscape scale, ignition location and burned area by forest fires are the result of a complex interaction between climatic factors, topography and land use (Moreira et al., 2011). Figure 1 illustrates the effect of a wildfire in the province of Valencia in 2012. Modeling and simulation of the final shape of the burned area is an important task as this gives insight on whether or not an ongoing wildfire will pose a potential hazard to human life and property as well as to how much will it contribute to greenhouse gasses and particulate matter in the atmosphere.

Most fire spread models consider the physics of the burning process and include approaches based on the physics of fire and in raster spread processes (Bin Suliman et al., 2010; Drissi, 2015; Muzy et al., 2008; Vorob'ov, 1996) Physicists use concepts and mathematical models based on partial differential equations and ordinary differential equations to describe fire spread process (Pastor et al., 2003). Parameters of those equations however do not correspond directly to the parameters intuitively identified as influencing the fire spread. For example, slope, wind speed, biomass, and relative humidity of the vegetation at the fire front have been found as key factors in fire dynamics in forest ecosystems.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Environmetrics* published by John Wiley & Sons Ltd.



**FIGURE 1** A wildfire in the Valencian Province, august 2012 (left) and its effect on landscape (right)

Modeling of fire spread in two dimensions can be done using vector and raster-based approaches. For vector based simulation of fire spread, it is necessary to simulate fire as a continuously expanding polygon (Anderson et al., 1982). This is the base of the FARSITE model (Fire Area Simulator, Finney & Andrews, 1994), of wide use in the area of fire ecology. However, the vector approach is highly dependent on the spatial scale used and requires the evaluation of a complex convex hull, which is computationally expensive.

An alternative approach is to use raster schemes, in which the domain of interest  $D$  is partitioned in square or hexagonal cells that form a lattice covering the domain. Fire spreads to neighboring cells using contact cell methods or using heath-accumulation methods (Berjak & Hearne, 2014; Clark et al., 1996; Frandsen & Andrews, 1979; Green et al., 1990; Hargrove et al., 2000; Kourtz & O'Regan, 1971; Yassemi et al., 2008).

Two main approaches to simulate fire spread in a rasterized domain are cell contact and heath accumulation (Green et al., 1983). The contact cell method has high computation efficiency. It interprets fire spread as a series of ignitions that span the area of an individual cell. (Frandsen & Andrews, 1979; Green et al., 1990; Kourtz & O'Regan, 1971) used this approach extensively, but their events are generated considering only the influence of the fastest spreading neighbor, neglecting that fire spread to a new unburned cell may be based upon the effect of multiple neighboring cells or prior heating (Green et al., 1983).

The heat accumulation approach considers the rate of spread of fire into a cell or pixel at a given time  $t$  as the sum of the contribution of ignited cells in its neighborhood during previous times. Despite its appealing to simulate fire spread, it is demanding computationally and requires the evaluation of some model parameters empirically (Yassemi et al., 2008).

In this article, we consider a random spread model, based in a raster contact approach. Fire ignitions are assumed to ignite at random within the interest domain  $D$  according to an underlying spatial point process  $\mathcal{X}$  with intensity function  $\lambda(x)$  and density

$$f(x) = \frac{\lambda(x)}{\int_D \lambda(u)du} ; x \in D. \quad (1)$$

The ignition point  $x$  lies inside a pixel or cell  $v$  whose state changes from “unburned” to “burned”. Note that this is equivalent to choose a pixel at random in  $D$  with probability proportional to  $f(x)$ . Conditional on an ignition occurring at pixel  $v$ , neighboring pixels burn with probability proportional to the intensity function. We show how the model works using data from two wildfires occurring in 2016 and compare model performance using receiver’s operation curve (ROC) and precision-recall (PRC) curve analysis. Simulations to test the model performance and to simulate the Artana and Catcaixent wildfires spread were done using a program code written in R-language.

The rest of the manuscript is organized as follows: In Section 2 we describe the model and used and provide all the details needed to clarify the methods used. In Section 3 we describe the data sets of this study. Section 4 presents the results and discussion with their implications in terms of computing efficiency in model fitting. Finally, Section 5 presents the conclusions, followed by references.

## 2 | METHODOLOGY

### 2.1 | Model for fire spread

To model the propagation of wildfires in the study area, we will use a raster approach. Let  $D \in \mathbb{R}^2$  be a geographic area of interest, partitioned in  $N$  sub-areas or pixels  $v_{it}; i = 1, \dots, n$ , observed at discrete but arbitrary times  $t_0, t_1, \dots, T$ . We will assume that at any time  $t$  each pixel belongs to one of the following mutually exclusive categories: Burned ( $B_t$ ), fire front ( $F_t$ ), neighboring the fire front ( $V_t$ ) and unburned ( $U_t$ ) (Figure 2). Note that  $D_t = B_t \cup F_t \cup V_t \cup U_t$  and that for  $t + 1$  some pixels in the complement of  $B_t \cup F_t$  may join  $F_{t+1}$ .

We assume that for each pixel there exist a probability probability of fire ignition  $p_i$  which depends on local conditions through a set of covariates  $Z_1, \dots, Z_p$ . This assumption allows to capture the spatial variability of fire probability often observed in real situations. For example, pixels covered by some vegetation types have higher fire probability compared to others (Boychuk et al., 2007, 2009; Díaz-Avalos et al., 2001).

Our interest is to model the transitions of pixels from  $V_t$  to  $F_t \cup B_{t+1}$ . We will assume a null probability that pixels in  $V_t$  at any time might caught fire from sources other than sparks and heat from contact with pixels in  $B_t$ .

In an analogous way to Markov Random Fields, we assume that each pixel is surrounded by a set of neighboring pixels  $N_i$ . This set of neighbors is defined in our case as the set of nearest neighbors of order  $k, k = 1, 2$ . Given that a fire has ignited at site  $s_i \in W$ , it spreads to its neighbors  $s_j \in V_i$  according to their fire probabilities  $p(s_j)$ . We assume that  $p(v) = 0, s \in B$  and that once a fire has ignited at a pixel  $v_i$  at time  $t$ ,  $v_i$  becomes a member of  $B$  for  $t + k, k = 1, \dots, K$ . The  $p(v_j)$  may be obtained from the intensity function  $\lambda(x_i)$  as is described in the next subsection.

### 2.2 | Estimation of the $p(v_i)$

To obtain the fire probabilities at each pixel in  $D$ , we first estimate the intensity function given the observed pattern of wildfires in  $D$  using methods for point spatial process modeling (Baddeley et al., 2016). Because we are considering only large sized fires, we fitted an area interaction point process model to the observed fire pattern (Baddeley & van Lieshout 1995). Area interaction point process models are useful to model spatial point patterns with moderate clustering to moderate repulsion (Aragó et al., 2016; Baddeley & van Lieshout, 1995). For any point pattern  $\mathbf{x} = \{x_1, \dots, x_n\} \in D$ , this class of models has a density function of the general form

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{A(\mathbf{x})}, \quad (2)$$

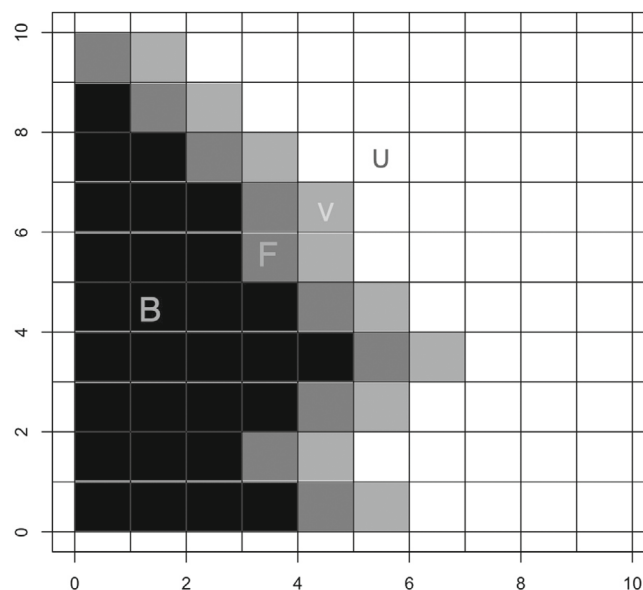


FIGURE 2 Composition of pixels during an ongoing wildfire

where  $A(\mathbf{x}) = [\cup_{i=1}^n B(x_i; r)] \cap \mathcal{D}$  is the set formed by the union of discs of radius  $r$  centered at the events  $x_i$  of the point process;  $\beta$ ,  $\gamma$  and  $r$  are model parameters and  $\alpha$  is the normalizing constant. The interaction parameter  $\gamma$  can be any positive number. If  $\gamma = 1$ , then the model reduces to a homogeneous Poisson process, if  $\gamma < 1$ , then the process is regular, while if  $\gamma > 1$  the process is clustered. Two points interact if the distance between them is less than  $r$ . For any finite point process, the Papangelou conditional intensity function is defined as (Coeurjolly et al., 2015)

$$\lambda(u|\mathbf{x}) = \frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})} \quad (3)$$

his is, for a given point pattern  $\mathbf{x}$ , the conditional intensity function for a new event at  $u \in \mathcal{D}$  is the conditional density function of  $u$  given the point pattern  $\mathbf{x}$ . Given a spatial point pattern  $\mathbf{x} = \{x_1, \dots, x_n\}$  with conditional intensity  $\lambda(u, |\mathbf{x})$ , parameter estimation and statistical inference for the intensity function is done using the pseudo-likelihood function,

$$PL(\theta, \mathbf{x}) = \left[ \prod_{x_i \in \mathbf{x}} \lambda_{\theta}^*(x_i|\mathbf{x}) \right] \exp \left[ - \int_{\mathcal{D}} \lambda_{\theta}^*(u|\mathbf{x}) du \right]$$

and therefore, the log pseudo-likelihood is

$$\log[PL(\theta, \mathbf{x})] = \sum_{i=1}^{n(\mathbf{x})} \log(\lambda_{\theta}^*(x_i|\mathbf{x})) - \int_{\mathcal{D}} \lambda_{\theta}^*(u|\mathbf{x}) du.$$

If the integral in the above equation is approximated by some quadrature rule, the approximated pseudo-likelihood becomes

$$\log[PL(\theta, \mathbf{x})] = \sum_{i=1}^{n(\mathbf{x})} \log(\lambda_{\theta}^*(x_i|\mathbf{x})) - \sum \lambda_{\theta}^*(u, |\mathbf{x}) w_j, \quad (4)$$

where  $u_j$  are quadrature points and  $w_j$  are quadrature weights. Details on the estimation method are beyond the scope of this work, but the interested reader may find a detailed technical description in (Diggle, 2021; Diggle et al., 1994).

We included Land Use, elevation, Slope, Aspect and Combustion Index as covariates in our model. Land use is a categorical variable with the coding shown in Table 1. The covariate values for each pixel and each wildfire were allocated to the entries of a matrix  $Z(x)$ . The covariate values for each wildfire in the data base were allocated using GIS software. The logarithm of the conditional intensity function for the Area-Interaction point process model takes the general form:

$$\log[\lambda(x|\mathbf{x})] = \mathbf{Z}(x)\beta + A_x(\mathbf{x})\gamma. \quad (5)$$

TABLE 1 Land use in Castellón and codes assigned

Code	Land use
20	Rainfed cropland
30	Irrigated cropland
70	Sparsely vegetated area
80	Vegetated low-lying areas on regularly flooded soil
90	Bare areas
100	Complex cropland
110	Mosaic cropland/natural vegetation
120	Mosaic of natural (herbaceous, shrub, tree) vegetation
200	Water bodies

Note: The categories represent an aggregation of a more diverse categorization for land use published by Heymann et al., 1994.

We point out that  $\mathbf{Z}(x)$  is a matrix with the values of covariates at location  $x$  in its columns,  $\beta$  is a vector with regression coefficients (Renner et al., 2015; van Lieshout, 2009). We consider as wildfires those forest fires with final size above 10 ha, which correspond to the 20th percentile of the empirical distribution of the wildfires recorded in the study area.

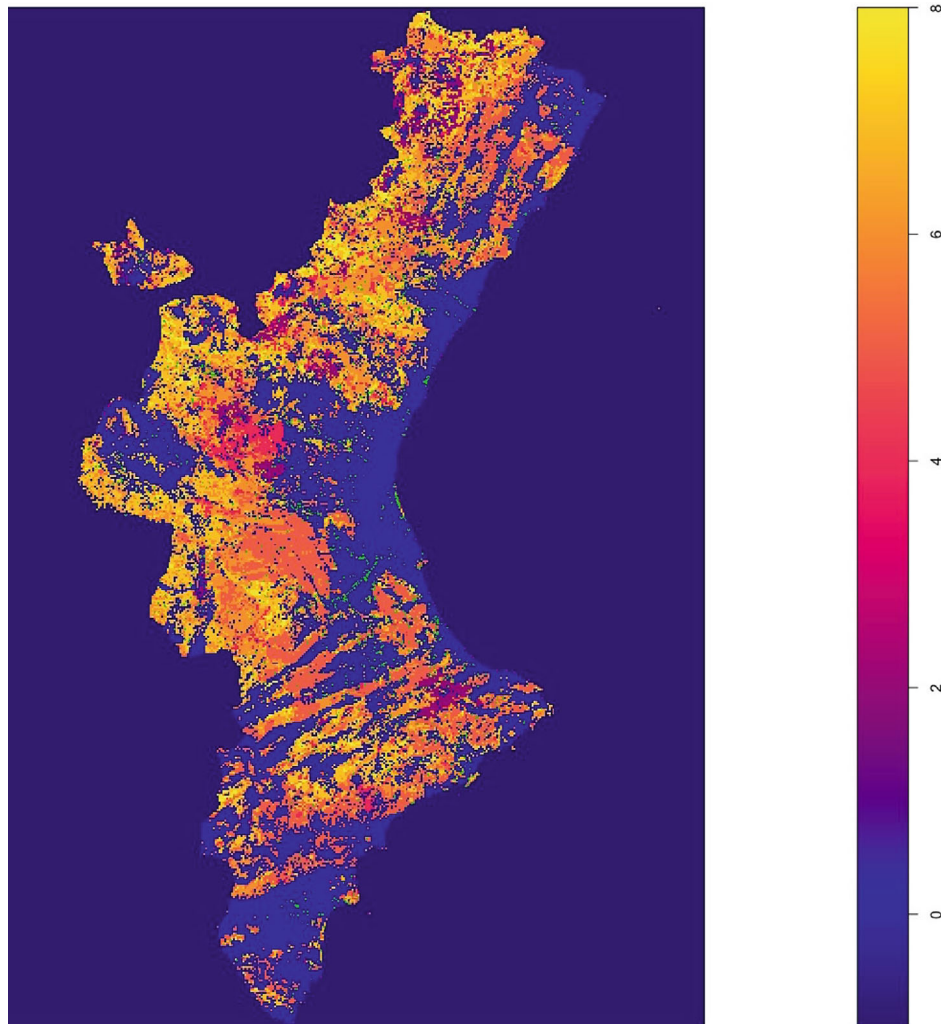
With the estimates of the intensity function at each pixel  $v_i, i = 1, \dots, N$  probabilities of fire occurrence at each pixel may be approximated using the standardization

$$p(v_i) = \frac{\lambda(v_i)dxdy}{\sum_{j=1}^N \lambda(v_j)dxdy}.$$

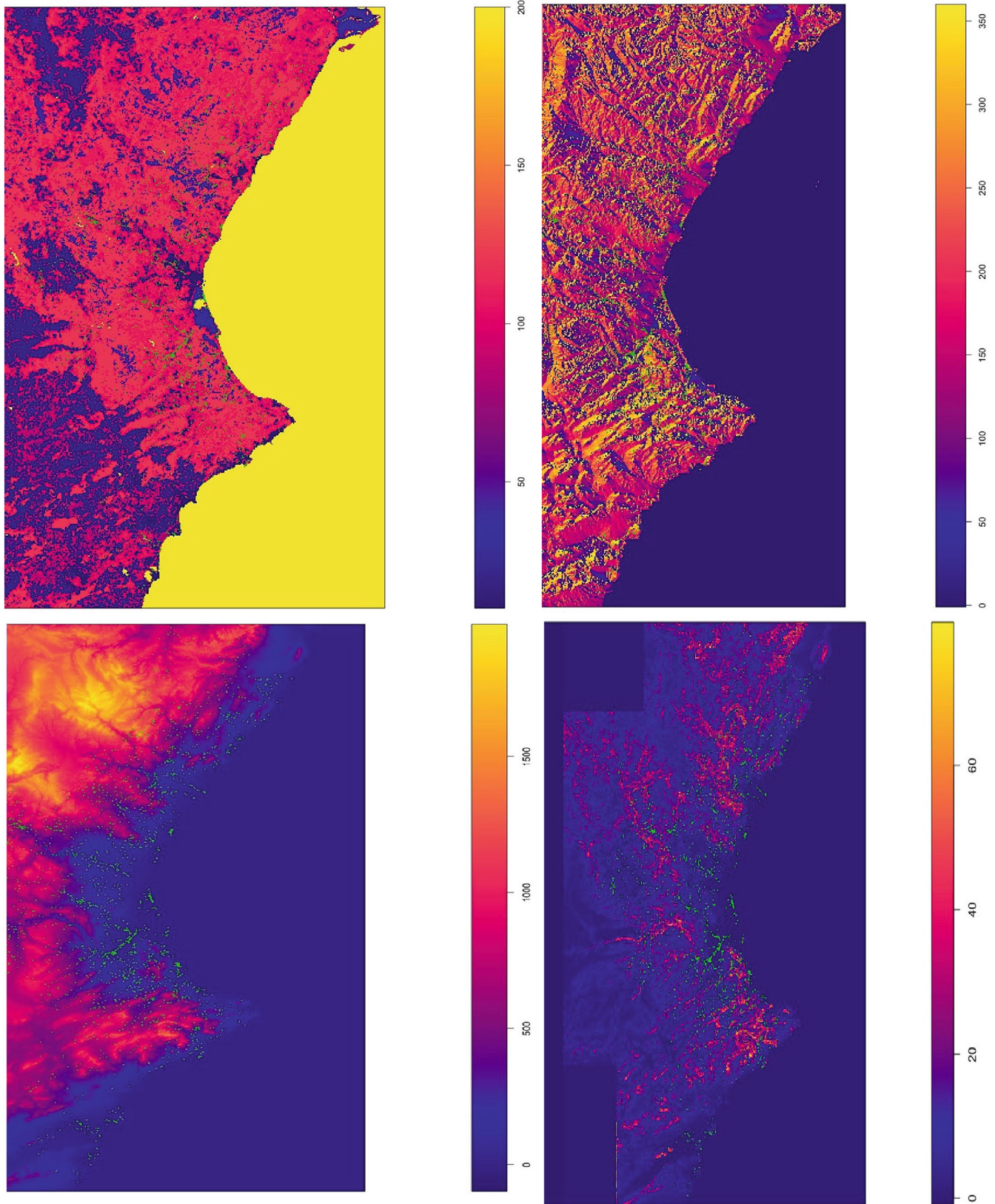
The area-interaction point process model was selected because its flexibility to capture different degrees of clustering and repulsion. The point process model for wildfires in Valencia included the covariates lithology, combustion index (Figure 3), land use, orientation, elevation and slope (Figure 4). Important covariates such as rainfall and wind speed and direction were not available for the study area and were not included in the model.

### 2.3 | Simulation of the fire spread process

We simulated the process of fire propagation using simulated ignitions as follows:



**FIGURE 3** Combustion Index (right) in the Valencian Community. The small green dots show the wildfires occurring in the Valencian Community between 1993 to 2015.



**FIGURE 4** Landuse, orientation, elevation and slope for the Valencian Community. The small green dots show the wildfires occurring in the Valencian Community 1993 to 2015.

1. Set  $t = 0$  and pick a pixel  $v$  at random in  $D$ , say  $v_{it}$
2. Sample a uniform random number

$$u \in (0, \max_{\{v: v \in D\}} f(v)).$$

3. If  $u < f(v_{i,t})$  then set  $t = t + 1$  and  $v_{i,t+1} \rightarrow F_{t+1}$
4. Identify the nonburning neighbors of  $F(t + 1)$  and repeat steps 2 and 3 for all of them and update the pixels in the fire front  $F_{t+1}$
5. Repeat step 4 until no extra pixels burn after  $K$  iterations or  $K = K_{\max}$ , where  $K_{\max}$  is the maximum number of iterations fixed by the user. In our case,  $K_{\max} = 200$

The simulation scheme described above corresponds to the so called *importance sampling*, modified so that only unburned pixels neighboring the fire front are candidates to become burned. In applications, the point of origin of a wildfire is known so step 1 above may be switched to pick the wildfire origin. In most cases, fire spread is also associated to ground wind speed, which induces some degree of anisotropy in the spread process. Such anisotropy may be incorporated in the simulation algorithm by giving more weight to pixels of  $V_i$  in the downwind direction of the burning area  $B_i \cup F_i$  (Boychuk et al., 2009). The above described algorithm was implemented using a code written in R-language.

To assess model performance we simulated random polygons using the following scheme:

1. Select 100 points of an ellipse centered at the origin pixel  $v$  in the previous simulation process with fixed main and second axes and orientation. The orientation of the resulting ellipse will be considered the anisotropy direction.
2. Add a random noise to the coordinates of the selected 100 points of the previous step and take a sample of 15 of those points. This number of points provided a nonregular polygon that resembled the shape of the burned area caused by wildfires.
3. Close the resulting polygon and use it as the simulated burned area.
4. Locate the pixel of origin of the simulated fire near the major axis of the random polygon, upwind to the wind direction to be simulated.

Note that the way we simulated the fires is independent of the fire spread model. We simulated four random polygons corresponding to the orientations SW-NE, S-N, SE-NW, and E-W, all using the same random centroid. For each polygon, we simulated 49 realizations of the fire spread model for the two possible wind directions (e.g., N and S for the S-N polygon, and so on) with maximum time  $t = 200$ . For each  $t \in \{1, \dots, 200\}$  the proportion of times a given pixel was classified as “burned” estimates of the probability of fire spreading to it from its neighbors. A detailed description of the simulation scheme is shown in Table 2.

**TABLE 2** Origin coordinates, anisotropy weights and wind directions selected for the simulations of the random spread model proposed, in a random polygons within the Valencian Community, Spain

Simulation No.	Origin		Anisotropic weights								Wind direction
	Long	Lat	N	W	S	E	NW	SW	NE	SE	
1	678,200	4,244,600	0.35	0.2	0.2	0.35	0.15	0.15	1.602	0.15	SW → NE
2	686,000	4,250,000	0.2	0.35	0.35	0.2	0.15	1.602	0.15	0.15	NE → SW
3	683,600	4,242,800	0.33	0.33	1.69	0.33	0.25	0.25	0.25	0.25	S → N
4	681,800	4,251,800	1.69	0.33	0.33	0.33	0.25	0.25	0.25	0.25	N → S
5	679,400	4,250,600	0.15	0.15	0.35	0.35	0.15	0.15	0.15	1.602	NW → SE
6	686,600	4,246,400	0.15	0.15	0.35	0.35	1.602	0.15	0.15	0.15	SE → NW
7	678,200	4,246,400	0.15	0.15	0.15	1.602	0.15	0.15	0.35	0.35	W → E
8	686,000	4,248,200	0.15	1.602	0.15	0.15	0.15	0.15	0.35	0.35	E → W

Note: The eight random polygons shared the same centroid.

**TABLE 3** Classification matrix obtained for a fixed probability threshold value used to classify pixels as burned or unburned after 49 simulations of the fire spread model with  $t = 200$

		Classified	
		Burned	as Unburned
True State	Burned	True positive ( $C_{11}$ )	False negative ( $C_{10}$ )
	Unburned	False positive ( $C_{01}$ )	True negative ( $C_{00}$ )

Using the estimated burning probabilities for each pixel, for a given threshold value, a pixel may be classified as burned if  $\pi_i > \tau$  and unburned otherwise. For the  $N$  pixels in the area of interest  $D$ , the classification process may be summarized in a classification or confusion matrix (Fawcett & Tom, 2006; Table 3):

With the fixed threshold the True Positive and False Positive rates (TPR or recall, and FPR, respectively) defined as (Egan, 1975)

$$\text{TPR} = \frac{C_{11}}{C_{11} + C_{10}}, \quad \text{FPR} = \frac{C_{01}}{C_{01} + C_{00}}.$$

Using those quantities we computed ROC curves and the area under the curve (AUC) statistic (Egan, 1975). We also computed the precision and recall statistics, defined as

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{and} \quad \text{Recall} = \frac{TP}{TP + FN},$$

where  $TP$  is the number of true positive (burned) pixels,  $FP$  is the number of pixels that were incorrectly classified as burned when they were not and  $FN$  is the number of pixels classified as nonburned when they actually burned. Precision is a measure of the proportion of burning pixels correctly identified, while recall is a measure of how many truly burning pixels are returned by the model. Precision-recall is useful to measure of success of binary prediction when there exists imbalance between the two classes. We present the results of this statistics in the form of precision-recall curves (PRC).

### 3 | RESULTS AND DISCUSSION

To estimate the probability of fire ignition in the whole Valencian Community we used the methods described in Section 2.2. The results of the Area-Interaction point process model fitted to the point pattern of fires occurring in the Valencian Community between 2006 and 2015 are shown in Table 4. Except for the coefficients associated to sparsely vegetated area and slope, the rest of the linear term coefficients in the model for the conditional intensity function were asymptotically significant. Compared to the base category (urban areas) where fires in vegetated areas are rare, the values for the conditional intensity function and hence the probability of a wildfire ignition, is high in the rest of the land use categories. The highest probability of ignition occurs in areas covered by vegetated low-lying areas on regularly flooded soil. Elevation and aspect have a small negative effect on the conditional intensity function. The combustion index on the other hand has a positive effect. The resulting probability of wildfire ignition for the Valencian Community is shown in Figure 5.

To compare how the model behaves under isotropy with spatial variation in the fire probabilities against constant fire probability, we carried out 49 simulations with the same pixel of origin, using the fire probabilities obtained from the intensity function fitted to the observed point pattern a  $67 \times 67$  pixel size square within the Valencian Community. The time interval for the spreading simulation was 200. The results are shown in Figure 6, where it may be seen that the shape of the burning probabilities under isotropy and spatial variation in the fire probabilities behave different when fire spreads in an environment with uniform fire probabilities. Thus, when the same anisotropic weights are used in the simulation process, fire spreads following the spatial pattern of the fire probabilities, as one would expect under the hypothesis that fire spreads depending on the fire risk of the pixels neighboring the fire front.

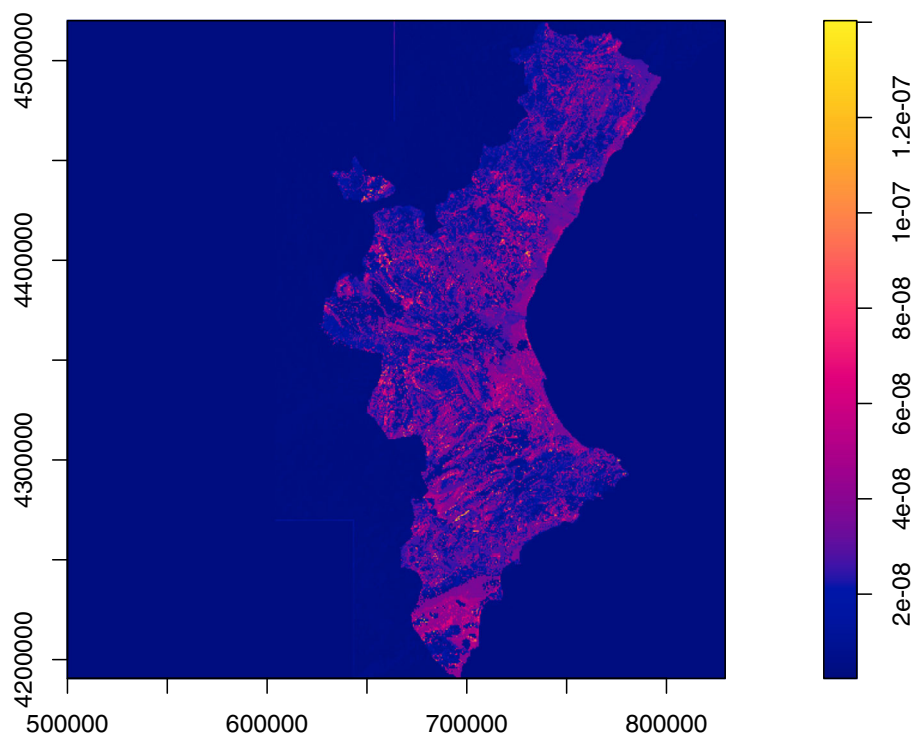
Figure 7 shows the results of the simulation of wildfire spreading on a random polygon within the Valencian Community using the anisotropy weights shown in Table 2. The time interval is  $[0, 200]$  and the number of realizations was 49. The figures represent the proportion of times a pixel was classified as burned at  $t = 200$ , which is equivalent to the



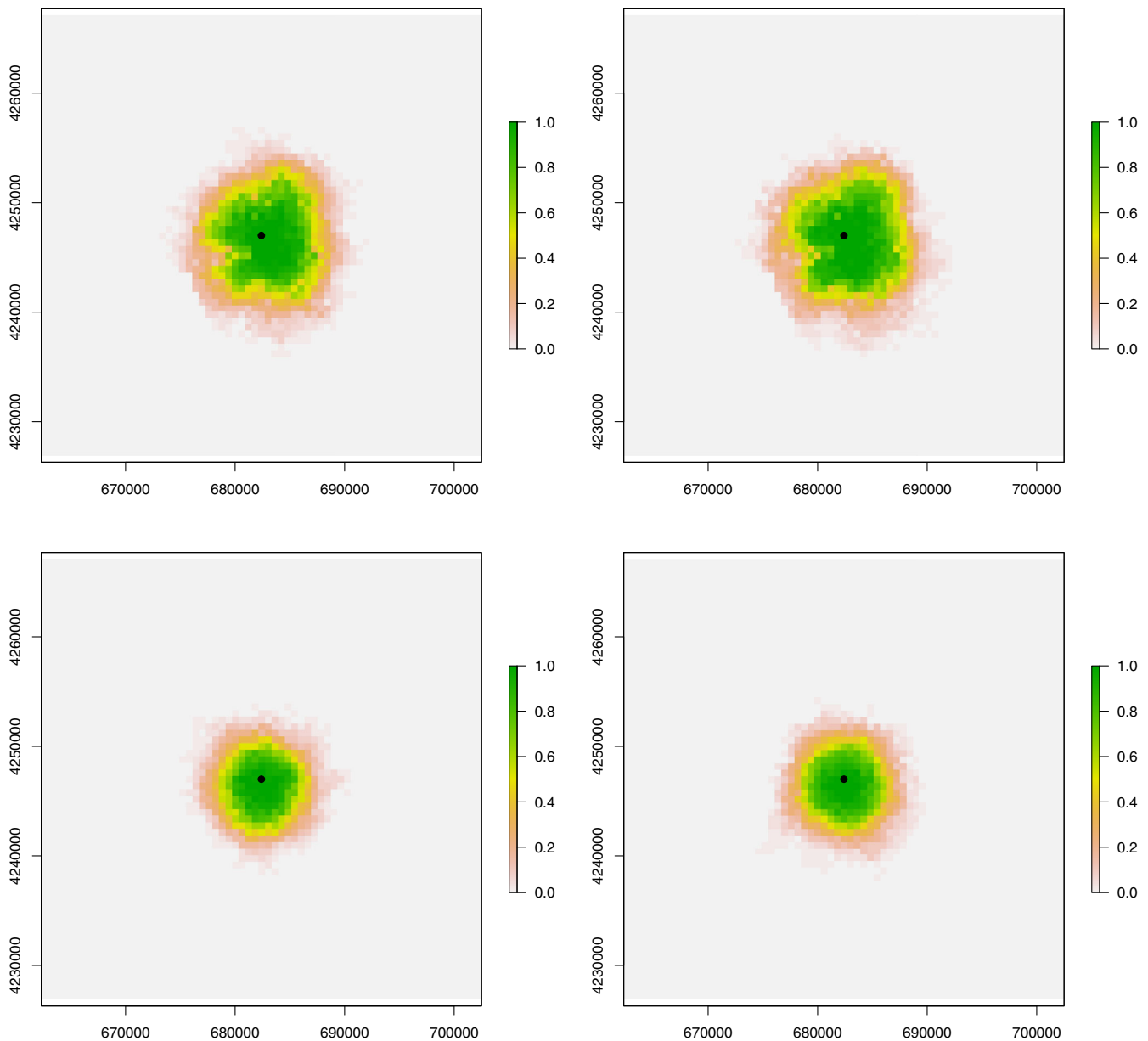
**TABLE 4** Point estimates and confidence intervals for the parameters of the Area-Interaction point process model fitted to the observed wildfires occurring in the Valencian Community between 2006 and 2015

Term	Estimate	S.E.	CI95.lo	CI95.hi	Ztest	Zval
Intercept	-17.900	0.088	-0.181	-17.756	***	-203.329
Rainfed cropland	-1.120	0.100	-1.320	-0.927	***	-11.232
Irrigated cropland	-0.478	0.163	-0.798	-0.157	**	-2.923
Forest	-0.598	0.113	-0.820	-0.376	***	-5.293
Sparsely vegetated area	-0.805	0.711	-2.200	0.587		-1.1329
Vegetated low-lying areas on regularly flooded soil	0.451	0.143	0.171	0.731	**	3.151
Bare areas	-0.629	0.264	-1.150	-0.112	*	-2.385
Complex cropland	-0.231	0.111	-0.449	-0.012	*	-2.073
Mosaic cropland/natural vegetation	-0.156	0.075	-0.305	-0.007	*	-2.060
Mosaic of natural (herbaceous, shrub, tree) vegetation	-0.382	0.076	-0.533	-0.231	***	-4.977
Water bodies	-3.120	0.249	-3.610	-2.632	***	-12.5441
Elevation	-5.67E-04	5.67E-05	-6.78E-04	-4.55E-04	***	-9.998
Slope	1.72E-03	2.10E-03	-2.39E-03	0.005		0.820
Aspect	2.30E-03	1.78E-04	1.95E-03	0.002	***	12.886
Combustion index	0.111	6.98E-03	0.097	0.124	***	15.854
Interaction	3.560	0.0631	3.440	3.685	***	56.468

Note: \*, \*\* and \*\*\* means: significance at the 10, 5 and 1%.



**FIGURE 5** Spatial distribution for the probability of wildfire ignition in the Valencian Community, Spain. The values in the map were obtained from an Area-Interaction spatial point process model, fitted to the wildfire occurrence pattern observed between 2006 and 2016.

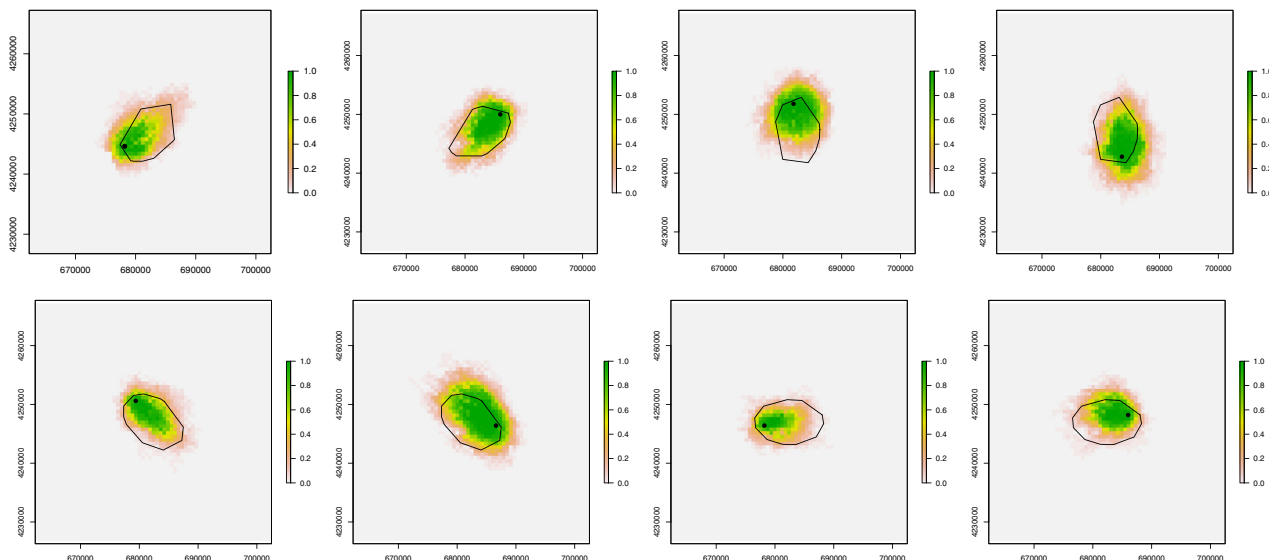


**FIGURE 6** Spread patterns of simulated wildfires in a  $67 \times 67$  pixel square in the Valencian Community using the same pixel of origin for the fire. The upper panel shows how fire would spread when fire probabilities are completely random in the square. The lower panel shows the pattern of fire spread under isotropy and constant fire probabilities in the square. In both cases the images show the burning probabilities obtained from 49 realizations of the fire spread model for a time interval from 1 to 200.

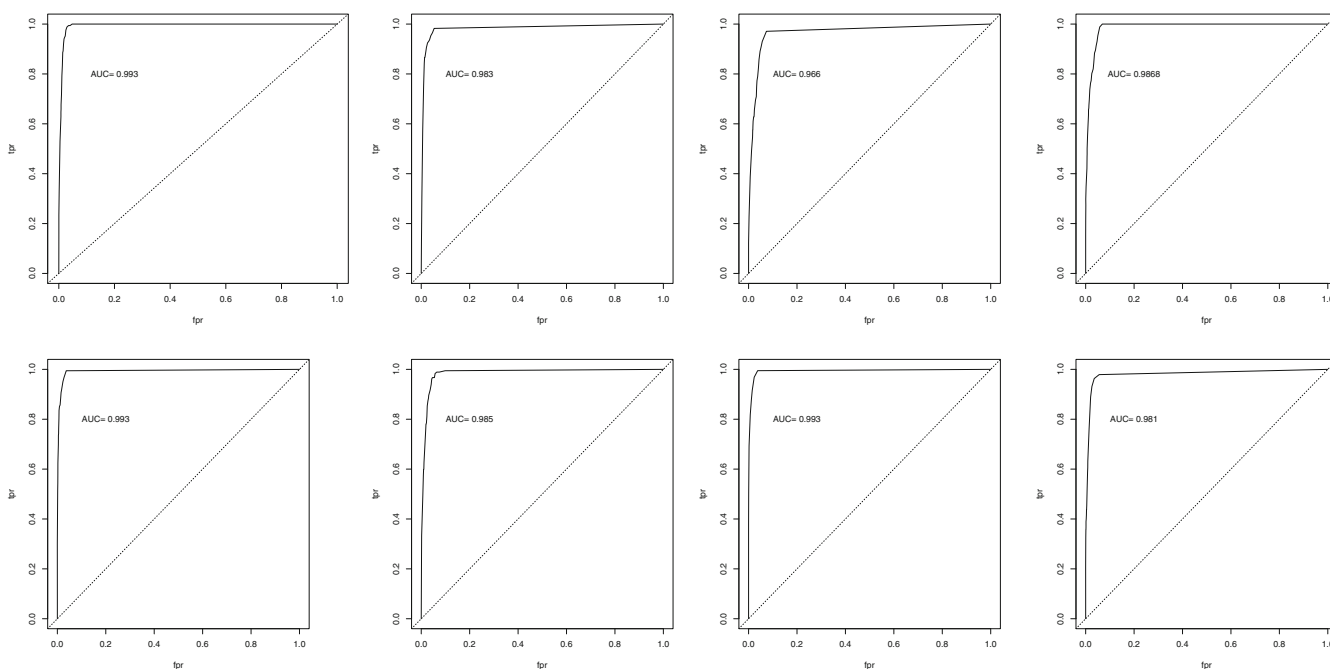
number of times the fire reached a given pixel. The different shapes and sizes of the simulated wildfires highlight the spatial variability of the burning probabilities in the Valencian Community, consequence of the spatial variation of the covariates incorporated in the model to estimate the intensity function of the point pattern of fires in the study area. The burned areas resulting from the simulations are irregularly shaped and different amongst them, following patterns similar to those observed for the burned areas by real wildfires.

The simulated fire spread patterns do not necessarily matched the anisotropy angle selected, because some of the pixels in the fire front become neighbors with a zone where fire risk is low, this is, where the intensity function of the point pattern used to compute the fire risk is very low. In those cases the fire spreading in the anisotropy direction slowed and the result was a burned area different to the polygon simulated as the fire shape.

In Figure 8 we show the ROC curves and the AUC statistic for the eight images shown in Figure 7. In all cases the empirical ROC curves are well above the 1:1 line, showing that the fire spread model proposed in this works acts well as



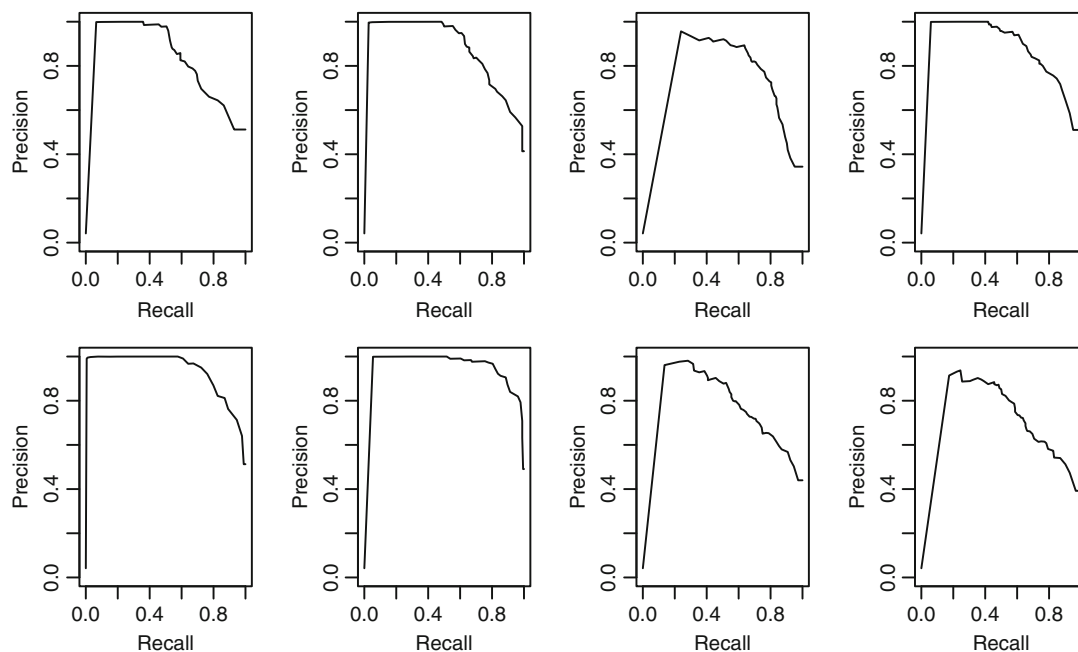
**FIGURE 7** Random polygons representing the burned area by wildfires (black lines) and burning probabilities obtained with 49 realizations of the fire spread model using the probabilities of fire obtained from the intensity function of the fire pattern observed in the Valencian Community between 2006 and 2016.



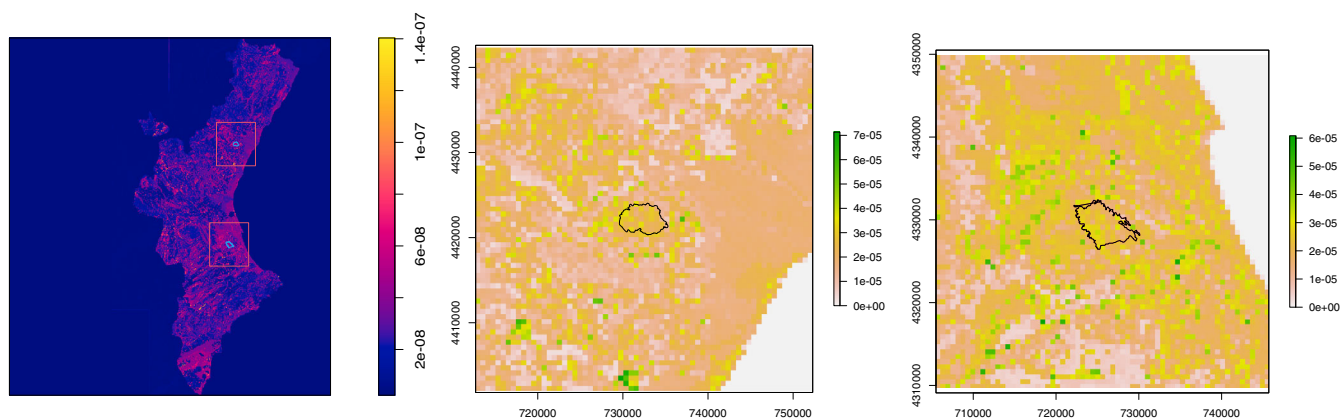
**FIGURE 8** ROC curves corresponding to the burning probabilities obtained with 49 realizations of the fire spread model using the probabilities of fire obtained from the intensity function of the fire pattern observed in the Valencian Community between 2006 and 2016.

a classifier of burned-nonburned pixels in the study area which corresponds random classification. In the context of the fire spread prediction the ROC curves as well as the AUC statistics show that the proposed fire spread model has a good classificatory performance and is a reliable predicting tool under the conditions of the simulation analysis.

ROC curves have been used widely to test the quality of binary classifiers in medicine and machine learning problems. Despite their usefulness, ROC curves are not the best option to assess the quality of binary classifiers when data show imbalance, as is our case given that we have a high proportion of unburned pixels in the study area. In Figure 9 we show the PRC curves for the 8 simulations used to test the model. The PRC curves show that if one is willing to accept a sensitivity



**FIGURE 9** PRC curves corresponding to the burning probabilities obtained with 49 realizations of the fire spread model using the probabilities of fire obtained from the intensity function of the fire pattern observed in the Valencian Community between 2006 and 2016.

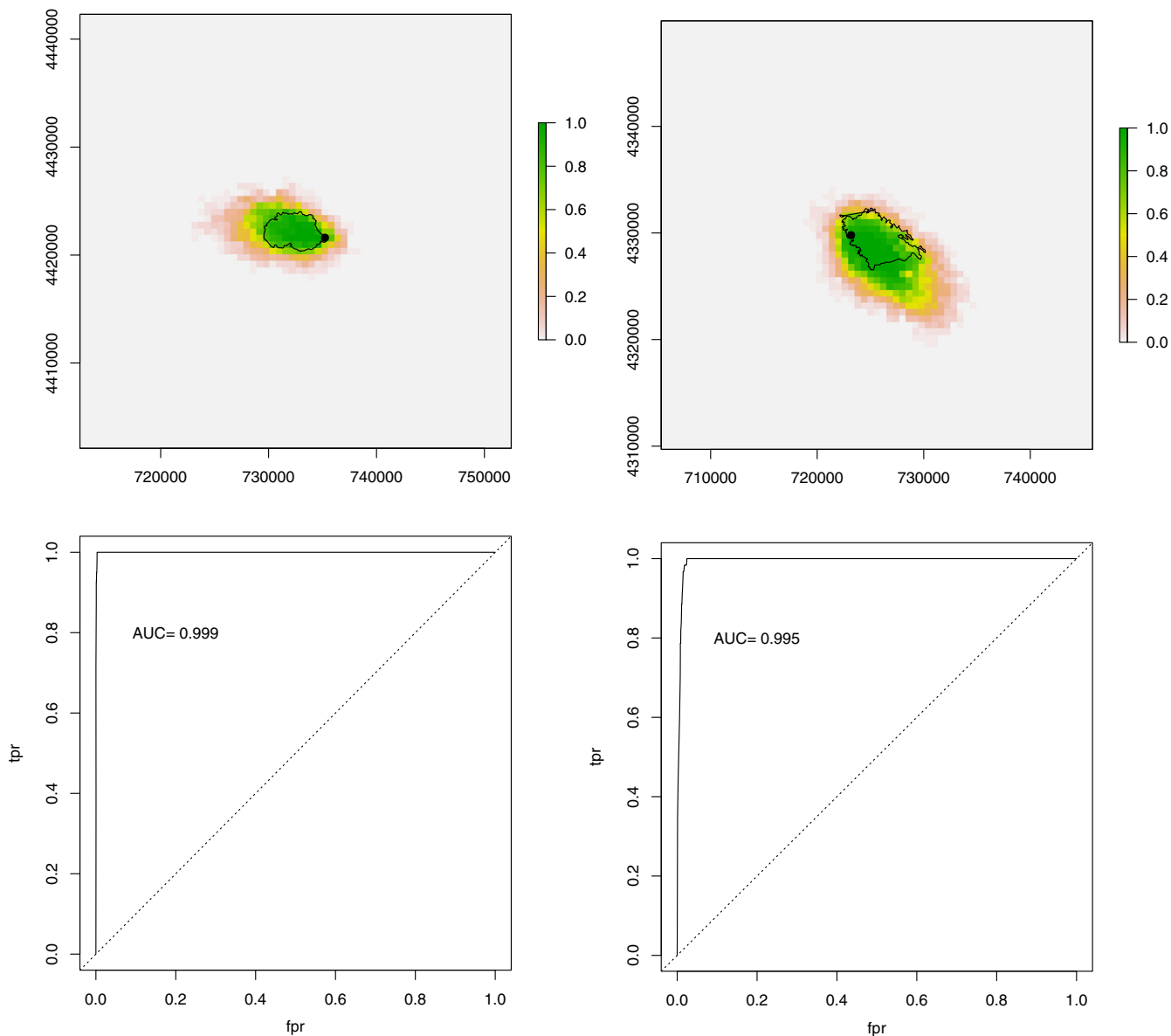


**FIGURE 10** Close up for the zone of the Valencian Community where the wildfire of Carcaixent occurred in 2016. The left plot shows the estimated fire probabilities obtained by fitting an area interaction point process model to wildfires between 2006 and 2016, as well as the squares used in the simulation process and the actual shape of the Artana (Northern square) and Carcaixent wildfires (Southern square). Center plot: close up of the square for the Artana region, showing probabilities of fire and the actual polygon of the 2016 wildfire. Right plot: close up of the square for the Carcaixent region, showing probabilities of fire and the actual polygon of the 2016 Carcaixent wildfire.

level of say, 80%, then the precision of the model we propose is above 80% in all cases. Thus, we may consider that the model shown here may be considered in the range of good to excellent classifier. This classification is consistent in terms of early retrieval, this is, the area under the curve the right quarter of the ROC curves (Saito & Rehmsmeier, 2015).

#### 4 | CASE STUDY: APPLICATION TO TWO WILDFIRES IN THE VALENCIAN COMMUNITY, SPAIN

During the last four decades, wildfires have become more frequent in the Mediterranean area of Europe (Moreira et al., 2011). The spatial pattern for wildfire occurrences in forests around the world has been related to external variables

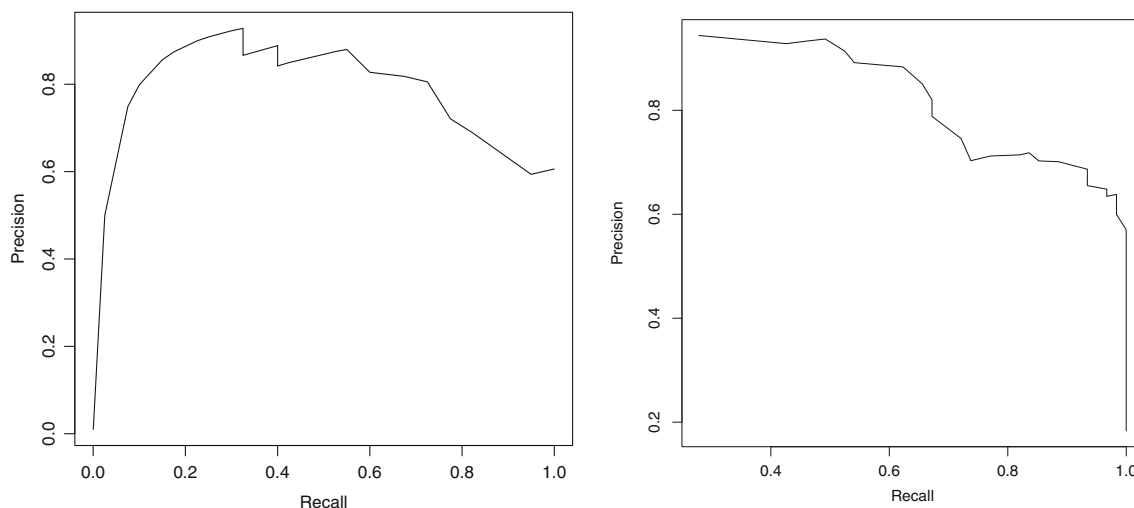


**FIGURE 11** Close up for the zone of the Valencian Community where the wildfire of Artana (left) and Carcaixent (right) wildfires that occurred in 2016. The upper plots correspond to the probabilities of fire for each pixel in the area after  $t = 200$ , using 49 simulations of the fire spread process and anisotropy in the main axis of the final polygon. The lower plots show the ROC curves and the AUC statistics for the corresponding upper image.

such as vegetation types, NDVI, local topography among others (Barros & Pereira, 2014; Carmo et al., 2011; Díaz-Avalos et al., 2016). Once a forest fire has been detected, it is of social interest to have a clue about the direction of its spreading and to have an estimate of the potential burned area produced by the wildfire. In this section we apply the fire spread model presented in Section 2 to a couple of large wildfires that occurred near Artana and Carcaixent, two small towns in the Valencian Community during 2016. There were more wildfire occurrences in 2016, but we choose these two because of the devastating effects they had in a relatively large area.

The Valencian Community (Figure 10, left), is a region particularly susceptible to wildland fires. Between the years 1986 to 2015, a total of 12,657 wildfires occurred in the region, burning 402,531 hectares (information from Prevención de Incendios Forestales, Valencian Government). During the mid-twentieth century, land use/cover change has been the main driver of the increase in the number of wildfires and the area affected by them.

In Figure 10 we show the resulting probabilities of fire (left), as well as the polygons of the Artana (in the north side) and Carcaixent (in the south) wildfires. The squares are the subsets of the Valencian Community used for the simulation



**FIGURE 12** Precision-recall plots for the wildfires of Artana and Carcaixent in 2016. The plots show that for sensibilities greater than 80%, the precision of the model proposed is well above 80%.

of the two wildfires and the estimated probabilities of fire inside each square. The plots at the center and right of Figure 10 are close ups of the squares around the Artana and Carcaixent 2016 wildfires. The true limits of the burned area for each wildfire are also shown with black lines. In the figure we may see that wildfires spread over areas with locally high probability of fire. The actual shape of the burned area however does not follow a circular-shaped pattern perhaps as a consequence of firefighting activities and factors such as winds and presence of fuel materials and different combustion index at each pixel.

In Figure 10 the upper panel of plots presents the burning probabilities obtained after 49 simulations of the fire spread model with  $t = 200$ . The corresponding ROC curves and AUC statistics are shown in the lower panel of the Figure 11. It can be seen that part of the predicted burned area (green) lies outside the actual perimeter of the Artana and Carcaixent wild fires. As mentioned earlier, this result in part due to the firefighting activities that block the limitless spread of a wildfire. However, notice that the final size of the simulated wildfires is bounded in all cases. The ROC curves for the simulated fire spread in Artana and Carcaixent show that the model has a good classificatory performance for the real data and we may say that the model proposed is adequate to describe the spreading pattern of these particular wildfires. This observation is also consistent with the PRC curves for the two wildfires shown in Figure 12. This process is similar to the one developed by (Boychuk et al., 2009), (Johnston et al., 2006), but in our case we do not rely on an autologistic approach.

The model proposed in this work has the advantage of being simple and easy to use. Incorporating factors that control fire spreading can be done by assigning appropriate weights for each direction. So if for instance wind speed increases in one direction at some section of the fire front, then that weight can be increased and run the model with the new weights. Also, if firefighting has been successful in some sections of the fire front, weights pointing to those direction can be modified accordingly.

The model has limitations as it does not include covariates explicitly. However, relevant covariates can be incorporated in the point process model fitted to obtain the fire risk at the pixel level, so covariate influence on fire spread is implicitly incorporated in the spatial fire risk estimation process. Although limited, the model is useful to obtain quick estimates of the potential burned area and to highlight areas where fire control efforts may decrease the rate of spreading of wildfires.

## 5 | CONCLUSIONS

We have presented a simple model and an algorithm to simulate wildfire growth and spread using spatial point process methods to obtain probabilities of ignition in different locations at a pixel level. The fire spread model can be seen as a Markov transition process in the sense that there is a state space (burned, unburned, fire front) and transition probabilities.

Albeit simplistic, the model provides a good approximation to the burned area and shape of the wildfires. In addition, the inclusion of covariates in the model improves the quality of the estimated intensity function used to obtain the local probabilities of fire used in the simulation and the detection of fire growth.

Finally, analysis and development of the random spread process methodology can be carried out in other fields in which the variables to be studied have a spatial growth in certain areas. The clearest example of the usefulness of this methodology presented here would be the study of epidemic spread in a geographic area. The inclusion of the variation in the study area of different covariates such as elevation, distance to fixed elements such as pollution sources, may improve the prediction of the epidemic spread.

## ACKNOWLEDGMENTS

We would like to thank the Environment Department of the Government of Valencia for the access to the digital map databases and Pau Aragó who has help in this access. Carlos Díaz-Avalos was supported by PAPIIT-DGAPA project number IG100221.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## ORCID

Pablo Juan  <https://orcid.org/0000-0002-2197-7502>

## REFERENCES

- Anderson, R. M., Gordon, D. M., Crawley, M. J., & Hassell, M. P. (1982). Variability in the abundance of animal and plant species. *Nature (London)*, *296*, 243–248.
- Aragó, P., Juan, P., Díaz-Avalos, C., & Salvador, P. (2016). Spatial point process modeling applied to the assessment of risk factors associated to forest wildfires incidence in Castellon, Spain. *European Journal of Forest Reserach*, *135*(3), 451–464.
- Baddeley, A., Rubak, E., & Turner, R. (2016). *Spatial point patterns. Methodology and applications with R*. Chapman and Hall.
- Baddeley, A., & van Lieshout, M. N. M. (1995). Area-interaction point processes. *Annals of the Institute of Statistical Mathematics*, *47*, 601–619.
- Barros, A. M. G., & Pereira, J. M. C. (2014). Wildfire selectivity for land cover type: Does size matter? *PlosOne*, *9*(1), e84760.
- Berjak, S., & Hearne, J. (2014). Spatial fire modeling in Mkuze game reserve: A case study. *ORiON*, *18*, 37–57. <https://doi.org/10.5784/18-0-184>
- Bin Suliman, M. D. H., Serra, J., & Mahmud, M. (2010). Prediction and simulation of Malaysian forest fires by random spread. *International Journal of Remote Sensing*, *31*(22), 6015–6032.
- Boyчук, B., Braun, W. J., Kulperger, R. J., Krougly, Z. L., & Stanford, D. A. (2007). A stochastic model for Forest fire growth. *INFOR: Information Systems and Operational Research*, *45*(1), 9–16.
- Boyчук, D., Braun, W. J., Kulperger, R. J., ... Stanford, D. A. (2009). A stochastic forest fire growth model. *Environmental and Ecological Statistics*, *16*, 133–151.
- Carmo, M., Moreira, F., Casimiro, P., & Vaz, P. (2011). Land use and topography influences on wildfire occurrences in northern Portugal. *Landscape and Urban Planning*, *100*, 169–176.
- Clark, J. S., Royall, P. D., & Chumbley, C. (1996). The role of fire during climate changes in an eastern deciduous forest at Devil's bathtub, New York. *Ecology*, *77*, 2148–2166.
- Coerjolly, J. F., Möller, J., & Waagepetersen, R. (2015). Conditioning in spatial point processes. hal-01241277v1.
- Díaz-Avalos, C., Juan, P., & Serra-Saurina, L. (2016). Modeling fire size of wildfires in Castellon (Spain), using spatiotemporal marked point processes. *Forest Ecology and Management*, *381*, 360–369.
- Díaz-Avalos, C., Peterson, D. L., Alvarado, E., Ferguson, S., & Besag, J. E. (2001). Space-time modelling of lightning-caused ignitions in the Blue Mountains, Oregon. *Canadian Journal of Forest Research*, *31*(2001), 1579–1593.
- Diggle, P. J. (2021). Conditional intensity: A powerful tool for modelling and analysing point process data. *Australian & New Zealand Journal of Statistics*, *63*(1), 83–92.
- Diggle, P. J., Fiksel, T., Grabarnik, P., Ogata, Y., Stoyan, D., & Tanemura, M. (1994). The estimation of the parameters beta and gamma is estimated by Maximun Pseudolikelihood. *International Statistical Review*, *62*(1), 99–117.
- Drissi, M. (2015). *Modeling the spreading of large-scale wildland fires*. In *Proceedings of wildland fires conference Missoula* (pp. 278–285). MT. Proc. RMRS-P-73. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Egan, J. P. (1975). *Signal detection theory and ROC analysis*. Academic Press.
- Finney, M. A., & Andrews, P. L. (1994). The FARSITE fire area simulator: Fire management applications and lessons of summer 1994. *Proceeding Interior West Fire Council Meeting*.
- Frandsen, W. H., & Andrews, P. L. (1979). *Fire behaviour in nonuniform fuels*. USDA For. Serv. Res. Pap. INT-232.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*, 861–874.
- Green, D. G., Gill, A. M., & Noble, I. R. (1983). Fire shapes and the adequacy of fire-spread models. *Ecological Modelling*, *20*, 33–45.

- Green, D. G., Tridgell, A., & Gill, A. (1990). Interactive simulation of bushfires in hetero-geneous fuels. *Mathematical and Computer Modelling*, 13(12), 57–66.
- Hargrove, W., Gardner, R., Turner, M., Romme, W., & Despain, D. (2000). Simulating fire patterns in heterogeneous landscapes. *Ecological Modelling*, 135(2–3), 243–263.
- Heymann, Y., Steenmans, C. H., Croissille, G., & Bossard, M. (1994). *CORINE land cover. Technical guide*. Office for Official. Publications of the European Communities.
- Johnston, P., Milne, G., & Kelso, J. (2006). *A heat transfer simulation model for wildfire spread*. In D. X. Viegas (Ed.), *V International conference on Forest fire research*. D. X. Viegas.
- Kourtz, P. H., & O'Regan, W. G. (1971). Model for a small forest fire to simulate burned and burning areas for use in a detection model. *Forest Science*, 17, 163–169.
- Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., & Koutsias, N. (2011). Landscape - wildfire interactions in southern Europe: Implications for landscape management. *Journal of Environmental Management*, 92, 2389–2402.
- Muzy, A., Nutaro, J., Zeigler, B. P., & Coquillard, P. (2008). Modeling and simulation of fire spreading through the activity tracking paradigm. *Ecological Modelling*, 219(1–2), 212–225.
- Pastor, E., Zarate, L., Planas, E., & Arnaldos, J. (2003). Mathematical models and calculation systems for the study of wildland fire behaviour. *Progress in Energy and Combustion Science*, 29, 139–153.
- Renner, I. W., Elith, J., Baddeley, A., Fithian, W., Hastie, T., Phillips, S. J., Popovic, G., & Warton, D. I. (2015). Point process models for presence-only analysis. *Methods in Ecology and Evolution*, 6, 366–379.
- Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- van Lieshout, M. N. M. (2009). *Moment analysis of the Delaunay tessellation field estimator*. In *Probability, networks and algorithms [PNA]*. CWI <https://ir.cwi.nl/pub/14465>
- Vorob'ov, O. Y. (1996). Random set models of fire spread. *Fire Technology*, 32, 137–173.
- Yassemi, S., Dragicevic, S., & Schmidt, M. (2008). Design and implementation of an integrated GIS-based cellular automata model to characterize Forest fire behaviour. *Ecological Modelling*, 210, 71–84.

**How to cite this article:** Díaz-Avalos, C., & Juan, P. (2022). Modeling the spatial evolution wildfires using random spread process. *Environmetrics*, e2774. <https://doi.org/10.1002/env.2774>