CSIRO PUBLISHING

www.publish.csiro.au/journals/ijwf *International Journal of Wildland Fire* **2009**, *18*, 1–11

Modeling and mapping wildfire ignition risk in Portugal

Filipe X. Catry[A,C], *Francisco C. Rego*[A], *Fernando L. Bacao*[B] *and Francisco Moreira*[A]

[A]Centre of Applied Ecology 'Prof. Baeta Neves', Institute of Agronomy, Technical University of Lisbon, Tapada da Ajuda, 1349-017 Lisbon, Portugal.

[B]Institute of Statistics and Information Management, New University of Lisbon, Campus de Campolide, 1070-312, Lisbon, Portugal.

[C]Corresponding author. Email: fcatry@isa.utl.pt

Abstract. Portugal has the highest density of wildfire ignitions among southern European countries. The ability to predict the spatial patterns of ignitions constitutes an important tool for managers, helping to improve the effectiveness of fire prevention, detection and firefighting resources allocation. In this study, we analyzed 127 490 ignitions that occurred in Portugal during a 5-year period. We used logistic regression models to predict the likelihood of ignition occurrence, using a set of potentially explanatory variables, and produced an ignition risk map for the Portuguese mainland. Results show that population density, human accessibility, land cover and elevation are important determinants of spatial distribution of fire ignitions. In this paper, we demonstrate that it is possible to predict the spatial patterns of ignitions at the national level with good accuracy and using a small number of easily obtainable variables, which can be useful in decision-making for wildfire management.

Additional keywords: geographic information systems, ignition occurrence, logistic regression, spatial patterns.

Introduction

Wildfires constitute a serious concern in many regions of the Mediterranean Basin, representing important social, environmental and economic impacts. Statistics show that both the burned area and number of fire ignitions increased in Portugal during the last decades, and in the period 2000–05, the average density of ignitions was three times higher in Portugal than in Spain, France, Italy and Greece combined (ENOC 2007**[AQ6]**). Since 2000, Portugal has registered an average of ∼28 500 fire ignitions every year (DGRF 2006). The increasing number of fire ignitions occurs despite more resources being allocated to vigilance, firefighting and prevention, including management plans, public education campaigns and the implementation of more restrictive legislation concerning human activities susceptible of causing wildfires.

The importance of knowing the spatial patterns of fire ignition is widely recognized, and the ignition risk (i.e. the chance of a fire starting as determined by the presence and activity of any causative agent, also defined as 'fire risk'; FAO 1986; NWCG 2006) is an essential element in analyzing and assessing fire danger (e.g. Johnson and Miyanishi 2001; Vasconcelos *et al*. 2001; Chuvieco *et al*. 2003; Bonazountas *et al*. 2005; Finney 2005; Roloff *et al*. 2005; Vasilakos *et al*. 2007). However, there are still several fire prevention plans that only consider the factors influencing fire spread and suppression difficulty (fuel, weather, topography, etc.). Although the high number of wildfire ignitions in Portugal can constitute a problem in terms of fire detection and firefighting resources dispersion, very few studies have concentrated on trying to characterize and understand the factors influencing their spatial occurrence and their causes. Until now, the official fire danger estimation systems and cartography being used gave little or no attention to specifically predicting the likelihood of a fire starting (IGP 2004*a*; Carreiras and Pereira 2006), which could be an important contribution to improving the predictive ability of these systems.

Although lightning is the primary cause of fire in several regions of the world (e.g. Rorig and Ferguson 1999), in other regions, including southern Europe, most contemporary wildfires are of human origin (e.g. Cardille *et al*. 2001; DGRF 2006; MMA 2007). For example, in Portugal and Spain, ∼97% of all investigated wildfires were human-caused (DGRF 2006; MMA 2007). A recent official report from the Portuguese Forest Services (DGRF 2006) shows that among the fires with known causes that occurred between 2000 and 2005 (∼3% of all fires), 49% were intentionally caused (arson), while 37% were due to negligence and 11% were accidentally caused.

Previous studies identified numerous factors influencing the spatial patterns of fire ignitions (Chou 1992; Vega-Garcia *et al*. 1996; Cardille *et al*. 2001; Cardille and Ventura 2001; Vasconcelos *et al*. 2001; Mercer and Prestemon 2005; Genton *et al*. 2006; Nunes and Duarte 2006; Catry *et al*. 2007*a*; Loboda and Csiszar 2007; Romero-Calcerrada *et al*. 2008). However, the effects of different factors on fire occurrence can vary among ecosystems and across spatial and temporal scales (e.g. Badia-Perpinyã and Pallares-Barbera 2006; Yang *et al*. 2007). Based on previous knowledge, we concentrated our analysis on factors related to human presence and activity, selecting a small group of easily obtainable variables that could potentially explain the spatial patterns of fire ignition in Portugal, namely: population density, distance to roads, land cover type and elevation. Our

 10.1071/WF07123 1049-8001/09/070001

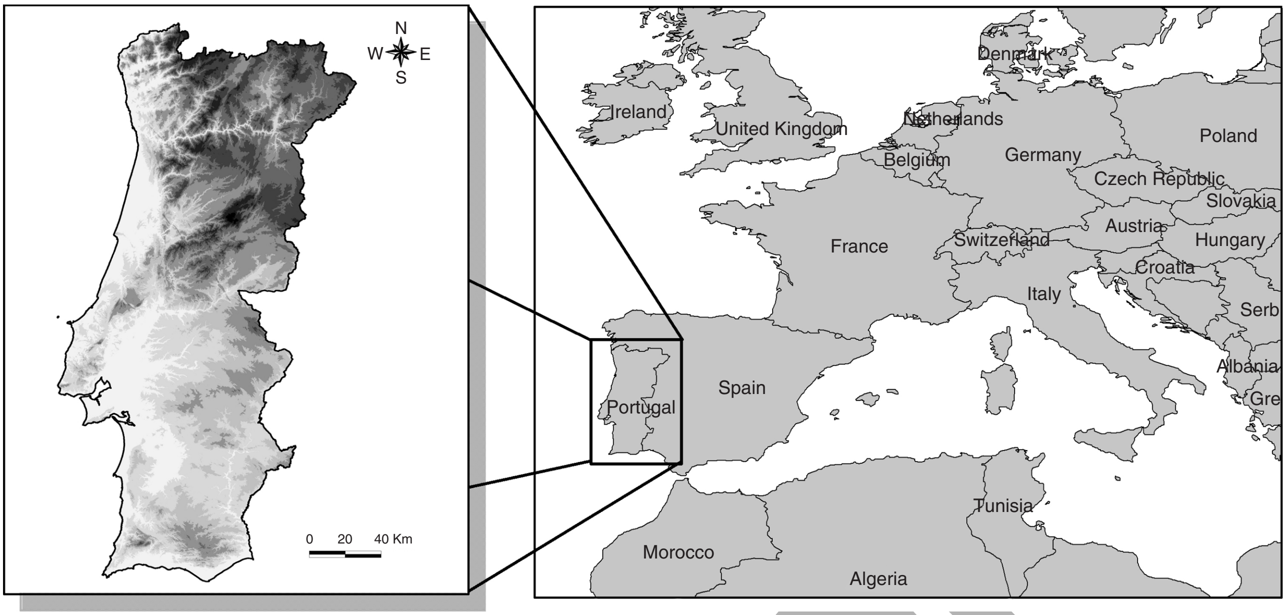

Fig. 1. General location of the Portuguese mainland (the study area) and country map (the darker the color, the higher the elevation**[AQ2]**).

general hypothesis was that the spatial distribution of ignitions is mainly determined by these factors. More specifically, we hypothesized that the areas with higher population density and closer to roads should have a higher probability of ignition (e.g. Cardille *et al*. 2001), because most fires in Portugal are human-caused and roads are major determinants of human access. Land cover was also hypothesized to be a determinant factor causing ignitions because different kinds of human activities (land uses) lead to different levels of risk, like those implicating a greater use of fire (e.g. traditional burning to eliminate agricultural residues), and because different land covers also have different fuel characteristics (type, load, moisture, flammability), which can also be a determinant for fire ignition (e.g. Yang *et al*. 2007). We also hypothesized that elevation could be positively related to ignitions because there are some human activities in mountain regions, such as pastoralism, which are known to cause frequent burns (renovation of pastures for livestock; DGF 2003), and lightning-caused ignitions also seem to be more common at higher elevations (e.g. Vazquez and Moreno 1998).

In the present paper, we aimed to build a parsimonious model to predict the spatial occurrence of fire ignitions in Portugal. We used a database including a layer with the location of 127 490 fire events that occurred during a 5-year period and a set of layers corresponding to potentially explanatory variables. We used a training subset to develop predictive models using logistic regression methods, and used a validation subset to evaluate the model's performance. Finally, we used one selected model and geographic information systems (GIS) techniques to produce an ignition risk map for the entire Portuguese mainland.

Methods

Study area

The study area constitutes the entire Portuguese mainland, which covers ∼90 000 km^2 in southern Europe (Fig. 1). Most of the country is included in the Mediterranean biogeographic region and there is a transition to the Atlantic region in the north. Mean annual temperatures range from ∼18°C in the south to 7°C at higher elevations in the north, and annual precipitation ranges from ∼400 to 2800 mm (IA 2003). The elevation ranges from sea level to 2000 m. Approximately 48% of the country area is used for agriculture, while forests and shrublands cover ∼27 and 19% respectively (DGF 2001). The population is estimated to be ∼10 million inhabitants, more concentrated in the north and centre coastal areas (INE 2003).

Model selection

Modeling fire occurrence has been carried out by several authors in different countries, using different methods and complexity levels, but logistic regression has been one of the most used (Chou 1992; Vega-Garcia *et al*. 1995; Cardille *et al*. 2001; Vasconcelos *et al*. 2001; Chuvieco *et al*. 2003; Preisler *et al*. 2004; Robin *et al*. 2006). Artificial neural networks (Vasconcelos *et al*. 2001; Chuvieco *et al*. 2003; Vasilakos *et al*. 2007) or classification and regression tree algorithms (Carreiras and Pereira 2006) have also been used to model fire occurrence. Neural networks are known to be more robust in modeling inconsistent or incomplete databases. However, as they are based on the use of hidden layers, it is difficult to find out which are the most significant variables affecting fire occurrence (Vasconcelos *et al*. 2001; Chuvieco *et al*. 2003). In the present study, as we were interested in determining the signs and significance of the variables affecting fire occurrence, we opted to use the logistic regression methods to model the ignition probability. Logistic regression is a useful method to predict the presence or absence of a given characteristic or event, based on the values of a group of predictive or explaining variables (Hosmer and Lemeshow 1989; Legendre and Legendre 1998). Additionally, logistic regression is quite flexible, in the sense that it accepts a mixture of continuous and categorical variables, as well as non-normally distributed ones (Hosmer and Lemeshow 1989; Legendre and Legendre 1998). Its analysis is based on the following function:

$$P = 1/(1 + e^{-z}), \tag{1}$$

where P is the probability of occurrence of the event, and z is obtained from a linear combination of the independent variables estimated from a maximum likelihood fitting:

$$z = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n,$$

where b_0 is the constant and b_n is the weighing factor of the variable x_n. The z values can be interpreted as a function of the probability of occurrence, and P converts z values in a continuous function (probability) that ranges from 0 to 1.

Fire ignition database

The dependent variable used in our analysis was the presence or absence of wildfire ignitions. For presence, we used the official wildfire database from the Portuguese Forest Services (DGRF), which contained the geographical coordinates and other characteristics (e.g. fire code, start and end date) of all fires that occurred in Portugal between 2001 and 2005. This database was then corrected, avoiding duplications, and removing all the records with inconsistent data, which included ignitions with erroneous coordinates or with missing dates. After these procedures, from an initial number of 137 204 ignition points, 127 490 remained in the database for analysis. As we also needed to account for non-ignition locations in order to model ignition risk, 191 235 points (1.5 times the number of ignition points) were randomly generated within the whole country, and were considered as non-ignition points. Ignition and non-ignition points were coded in a numeric binary format (1 – presence, 0 – absence), constituting the fire ignition dependent variable.

Because an evaluation of model performance (predictive ability) based on the same dataset used to build the model will probably result in an over-fit (Hosmer and Lemeshow 1989), we prepared two separate datasets. From the fire ignition database, we randomly selected 63 745 ignitions (50%) and 127 490 non-ignition points to build the model, constituting the training subset; we opted to use two times more non-ignition points (average density 1.5 points km^{-2}), to better represent the spatial heterogeneity of the country, as it is expected that its variability is larger than that found in the ignition sample. For the testing phase, we reserved the remaining fire ignitions and an equal number of non-ignition points, constituting the validation subset.

Geographic constraints

The majority of fire ignition coordinates in Portugal are associated with the nearest toponymic location, meaning that they do not have totally accurate coordinates, which is also a relatively common problem in other countries (e.g. Amatulli *et al.* 2007). In order to have an idea of the geographic inaccuracies associated with this location method, we used random points and verified that 96% of them were located at less than 500 m from the nearest toponimy; thus, we can estimate that 96% of fire ignitions would have a maximum error of 500 m, if they occurred randomly in the territory. However, the real errors in the fire database are probably considerably lower, because the majority of ignitions occur in areas where the toponymic density is higher. In fact, considering the seven districts with a higher number of fire ignitions (comprising 77% of all occurrences), we estimated that the majority of them (~80%) would have a maximum error of ~250 m. Although we believe that these positional uncertainties do not have a significant influence on our analysis, because we are mainly dealing with geographic information at relatively large spatial scales and with a very large sample, they influenced some subsequent decisions concerning the spatial resolution of cartographic layers (see next section).

Cartographic information

All the spatial analysis and cartographic production were made using GIS, mainly *ArcGIS* (ESRI 2005). Because of the mentioned ignition location uncertainties, and in order to reduce their potential influence on the analysis, the base raster maps were produced with 250-m spatial resolution, as it is recognized that an increment in cell size (i.e. generalization) can reduce or eliminate location accuracy problems (e.g. Koutsias *et al.* 2004). Next, we describe the base cartography used and the pre-processing steps performed to obtain the explanatory variable maps.

(i) Population density – this map was obtained using a database from the National Statistics Institute concerning the 2001 Census (INE 2003), which included the number of persons present in each parish. This information was assigned to the official parish map (IGP 2004*b*; vector format) and the population density (number of persons per km^2) was calculated for each of the 4050 parishes in the country mainland.

(ii) Land cover – we used the Corine Land Cover 2000 cartography at 1 : 100 000 scale (IA 2005; vector format). The Portuguese cartography identifies 42 land cover classes, which were grouped into six major classes according to their main characteristics: (1) agriculture (agricultural areas; covering ~47.8% of the country surface); (2) forests (covering 27.3%); (3) shrublands (covering 18.8%; including some natural grasslands); (4) urban–rural interspersed areas (covering 2.7%; including urbanized and other artificially modified areas); (5) sparsely vegetated areas (covering 1.9%); and (6) wetlands (covering 1.5%; also including water bodies, covering 0.1%). Class 4 corresponds mainly to areas where urban structures and other artificially modified areas are interspersed with other land uses (mainly agriculture); continuous urban fabric represents less than 6% of the total area of this class.

(iii) Distance to roads – this map was obtained using the Portuguese Itinerary Military Map in vector format at 1 : 500 000 scale (IGEOE 2005), representing the main national and regional roads. Distance (in m) from each location of the territory to the nearest road was calculated, producing a raster map with 100-m spatial resolution that was resampled to 250 m using the bilinear interpolation method (ESRI 2005).

(iv) Elevation – we used a digital elevation model (m) in raster format and with 90-m spatial resolution (NASA *et al.* 2004). This map was submitted to several operations, including projection transformation to be consistent with other data layers, and correction of negative and no-data values. Finally the map was also resampled to 250 m using the bilinear interpolation method.

Fire ignition characterization and modeling

The fire ignition database, including both fire ignitions and non-ignitions, was transformed into a vector point map and overlayed with all the other maps in order to gather all the information in a single database, where each presence–absence record contained the information of all the other layers. As we noticed the existence of a residual number of ignitions ($\sim$0.08%) located in dams or other water bodies (wetlands class) that were considered location errors, we opted to assign these points to the nearest land cover class.

In a first step, we performed a frequency analysis to characterize the spatial occurrence of fire ignitions. For that purpose, continuous variables were classified in intervals, and expected *v.* observed frequencies were registered and compared. Observed frequencies were the number of fire ignitions that occurred in the 5-year period in each class interval, and the expected frequencies were represented by the area of each class in the whole country, assuming that this would correspond to a random distribution. Comparisons between observed and expected frequencies were based on χ^2 statistics (Sokal and Rohlg 1987), using a significance level of 0.001. The same method was used by other authors (Badia-Perpinyã and Pallares-Barbera 2006).

In a second step, using the training dataset, we used logistic regression to fit models to predict the probability of ignition occurrence. As the proportion of ignitions was determined by the number of non-ignition localities that were defined, the obtained logistic model probabilities represent relative probabilities, rather than real ignition probabilities.

Continuous variables were very skewed; thus they were $\log(x + 1)$ transformed to approach normality and reduce variance. As models with transformed variables systematically showed a better fit to the data, they were retained, replacing the original variables. Additionally, the correlation between variables was analyzed using the Pearson correlation coefficient. The most correlated pairs of variables were distance to roads and population density ($r = -0.328$; $P < 0.01$), and population density and elevation ($r = -0.315$; $P < 0.01$); thus all the variables were used as candidates for model selection. For the multivariate logistic regression, the independent variables were selected using forward stepwise selection (forward likelihood ratio).

The results of the statistical tests performed should always be seen as an indication, as observations are bound to have some degree of temporal and spatial dependence. All analyses were carried out using *SPSS* software (SPSS 2006).

Production of the fire ignition risk map

In order to spatialize the model obtained in the multivariate logistic analysis, we had to prepare all the cartographic layers using GIS techniques. All the maps needed to be in raster format and coregistered in a common base 250-m cell size. The maps in vector format (population density and land cover) were directly converted to the raster format, while distance to roads and elevation maps were previously resampled using a bilinear interpolation technique. Following these steps, the model equation was spatialized through map algebra operations in a GIS environment, in order to obtain the ignition risk map. The map produced was classified into six risk classes, representing the different probabilities that a sample point corresponds to a wildfire ignition: extremely low (0–10%), very low (10–20%), low (20–40%), medium (40–60%), high (60–80%) and very high (80–100%).

Evaluation of model and ignition risk map performance

The assessment of the model's performance and adjustment were done by means of different standard approaches for logistic regression. First, signs of estimated parameters were checked to make sure they agreed with theoretical expectations based on previous knowledge of fire occurrence. The significance of each single variable was evaluated using the Wald test (Legendre and Legendre 1998), considering that the parameter was useful to the model if the significance level was lower than 0.001. The overall significance of the models was assessed through the Hosmer and Lemeshow goodness-of-fit test, which is a measure of how well the model performs (Hosmer and Lemeshow 1989; SPSS 2006). If the significance of the test is small (i.e. less than 0.05), then the model does not adequately fit the data (Norusis 2002; SPSS 2006).

Quantification of the model predictive ability was also done by comparing observed with predicted probability of ignition, both in the training and in the validation datasets. A confusion matrix (Congalton 1991) was used to assess the model classification accuracy, as a 2×2 classification table of observed *v.* predicted values. For this purpose, we had to establish some probability at which to accept the occurrence of an ignition. Cut-off points are used to convert probability of ignition to dichotomous 0–1 data, where cells with values below the cut-off are considered as non-ignition sites, while all above become predicted as ignition sites. Although in some cases a value of 0.5 is used, this threshold can be modified, and ultimately depends on the objectives of the user (Vega-Garcia *et al*. 1995; Vasconcelos *et al*. 2001; Chuvieco *et al*. 2003). The training dataset was used to construct classification tables for different cut-off points, helping define an optimal value. Two statistics were computed for each cut-off point considered: sensitivity and specificity. Sensitivity is the proportion of true positives that are predicted as events and specificity is the proportion of true negatives that are predicted as non-events. The optimal cut-off point corresponds to the value where both sensitivity and specificity reach the same proportion (e.g. Vega-Garcia *et al*. 1995; Vasconcelos *et al*. 2001), which in our case was 0.34.

Another procedure used to evaluate how well a model is parameterized and calibrated in presence–absence models is the ROC (receiver operating characteristic) analysis (Swets 1988; Pearce and Ferrier 2000; SPSS 2006). The ROC method has advantages in assessing model performance in a threshold-independent fashion, being independent of prevalence (e.g. Manel *et al*. 2001). The curve is obtained by plotting sensitivity *v.* specificity for varying probability thresholds. Good model performance is characterized by a curve that maximizes sensitivity for low values of specificity (i.e. large areas under the curve, (AUC)). Usually AUC values of 0.5–0.7 are taken to indicate

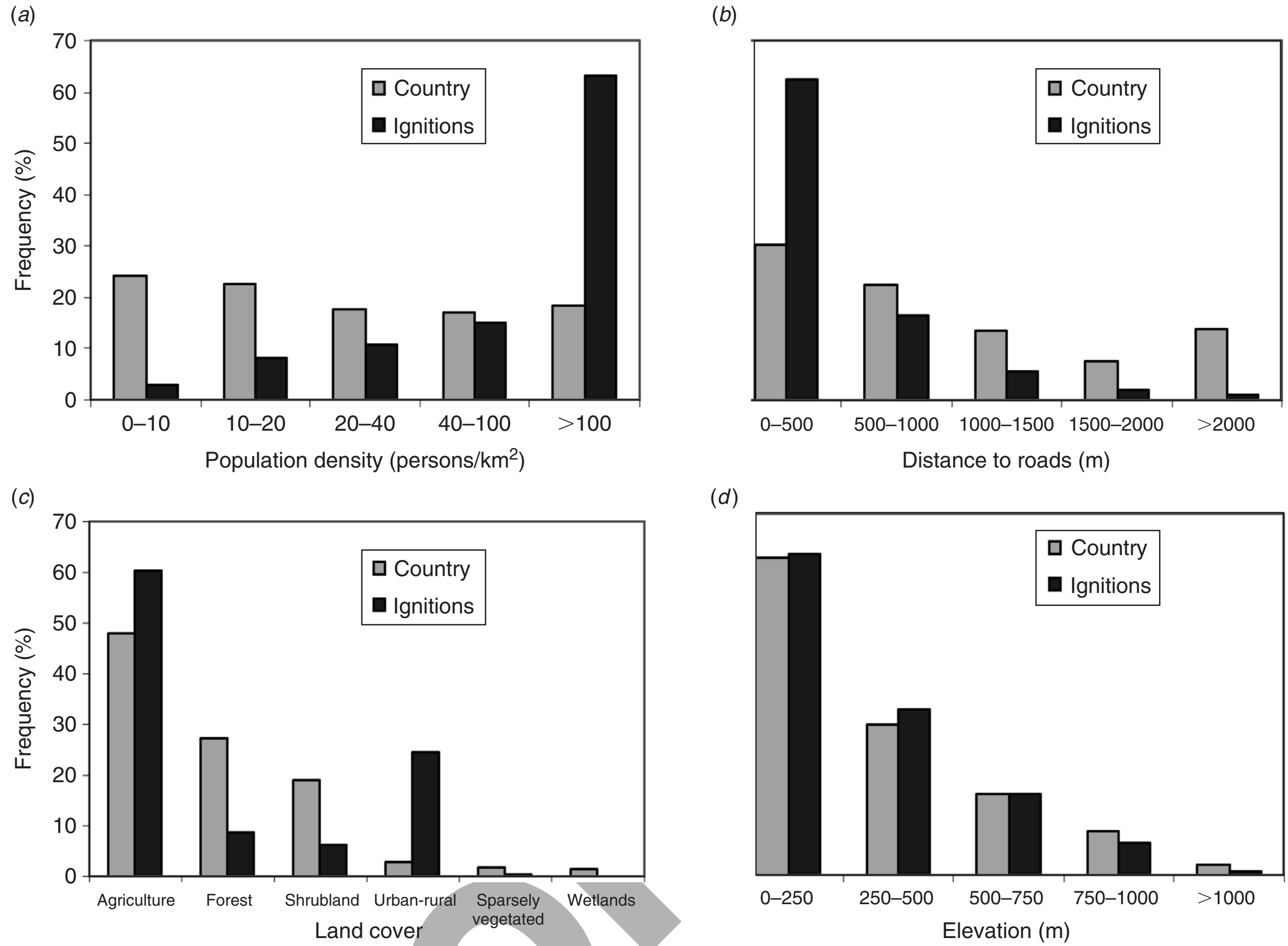

Fig. 2. Fire ignition frequency in relation to different variables: (*a*) population density; (*b*) distance to roads; (*c*) land cover; and (*d*) elevation. The expected frequencies (country) are represented by the area of each class in the country, assuming that this would correspond to a random distribution**[AQ3]**.

low accuracy, values of 0.7–0.9 indicate useful applications and values above 0.9 indicate high accuracy (Swets 1988).

The produced ignition risk map was also evaluated to assess its ability to predict ignition occurrence. This was also made through a confusion matrix using the validation dataset composed of 63 745 known ignition points, and with an equal number of non-ignition points, in order to evaluate the effect of model spatialization. Additionally, for each ignition point, we recorded the respective probability value present in the map, and analyzed the results by comparing the percentage of area covered by each risk class in the whole country with the percentage of ignitions occurring at each one. Ignition density in each risk class was also evaluated using all ignition points. At a final step, a regression equation was developed to predict average density of ignitions from the original explanatory variables.

Results

Frequencies of fire ignitions

During the period analyzed, the year 2005 registered the highest number of ignitions (27.2%), followed by 2001 (20.5%), 2002 (20.0%), 2004 (16.3%) and 2003 (16.1%). The majority of fire ignitions occurred between June and September (76.2% of the total), which is the fire season in Portugal. A preliminary analysis of the ignition spatial distribution in relation to the selected variables showed that ~98% of ignitions occurred less than 2 km from the nearest road, and regarding elevation, 98% occurred below 1000 m. Approximately 60% of ignitions were located in agricultural areas, and 25% occurred in urban–rural interspersed areas. Forested areas registered 8.5% of all ignitions, and uncultivated areas (including shrublands and natural grasslands) registered 6.2%. Concerning population density, although municipalities with more than 100 persons per km^2 only represent 21% of the territory, they registered 70% of all fire ignitions.

Comparisons between observed and expected frequencies of fire ignitions in relation to the selected variables are presented in Fig. 2. We used a homogeneity test to evaluate the differences between observed and expected frequencies, and confirmed the existence of significant differences in their spatial distribution in relation to all variables. The χ^2 homogeneity test for population density with respect to frequency of ignitions per ha of land showed that fire ignitions are more likely to occur in more populated areas ($\chi^2 = 1\,77\,138.8$, $P < 0.001$). The frequency

Table 1. Results of multivariate logistic regression model using four variables

Variables are ordered by decreasing importance. Full model goodness-of-fit statistic = 78 885.53 (d.f. = 8, $P < 0.001$); AUC (area under the curve) = 0.869 ± 0.001 ($P < 0.001$)

Variables	Coefficient	s.e.	Wald χ^2	d.f.	*P* value
Population density	0.820	0.005	24 266.36	1	<0.001
Land cover			9667.60	5	<0.001
Urban–rural	2.455	0.158	242.06	1	<0.001
Agriculture	1.672	0.157	113.89	1	<0.001
Shrublands	0.439	0.158	7.74	1	0.005
Sparsely vegetated	0.426	0.170	6.33	1	0.012
Forest	0.388	0.157	6.08	1	0.014
Elevation	0.585	0.007	6981.76	1	<0.001
Distance to roads	−0.166	0.002	5919.22	1	<0.001
Constant	−7.833	0.162	2336.07	1	<0.001

Table 2. Results of multivariate logistic regression model using three variables

Variables are ordered by decreasing importance. Full model goodness-of-fit statistic = 1688.19 (d.f. = 8, $P < 0.001$); AUC = 0.847 ± 0.001 ($P < 0.001$)

Variables	Coefficient	s.e.	Wald χ^2	d.f.	*P* value
Population density	0.875	0.005	30 968.93	1	<0.001
Distance to roads	−0.214	0.002	11 072.16	1	<0.001
Elevation	0.473	0.006	5492.23	1	<0.001
Constant	−5.890	0.049	14 552.78	1	<0.001

of ignitions also depended on the variable distance to roads ($\chi^2 = 82\,308.2$, $P < 0.001$), showing that ignitions are more likely to occur closer to the main roads. The land cover variable also influenced the occurrence of fire ignitions ($\chi^2 = 2\,57\,366.0$, $P < 0.001$), and all classes were significant. The class designated as urban–rural had nine times more ignitions than would be expected if they occurred randomly in the territory, and as expected, the opposite situation was observed in wetlands, where ignition frequency was nine times lower. Forested and uncultivated areas (including shrublands and natural grasslands) registered approximately three times less fire ignitions than if they were randomly distributed. Finally, ignitions were also significantly affected by elevation ($\chi^2 = 1906.2$, $P < 0.001$), but the differences between observed and expected frequencies are much less obvious than with the other three variables.

Fire ignition models

The obtained logistic regression model (Table 1) showed that the most influential variable explaining the spatial patterns of ignitions was population density, followed by land cover type, elevation, and distance to roads. The Hosmer and Lemeshow goodness-of-fit test showed adequate fit of the model to the data ($\chi^2 = 78\,885.53$, $P < 0.001$). The AUC for this model was 0.869, and its global accuracy was 79.8% (using the training dataset), both of which indicate good model adjustment.

The model obtained is represented by the following equation:

$$P_1 = 1/(1 + e^{-(-7.833+0.820Pop\_D-0.166D\_roads+0.585Elev+2.455Urb+1.672Agr+0.388For+0.439Shr+0.426Spa)}),$$

where P_1 is the probability that a point corresponds to a fire ignition, *Pop_D* is the population density (persons km^{-2}), *D_Roads* is the distance to the nearest road (m), *Elev* is the elevation (m), *Urb* is the land cover class representing urban–rural interspersed areas, *Agr* represents agriculture, *For* represents forest, *Shr* represents shrublands, and *Spa* represents sparsely vegetated areas (all variables but land cover are $\log(x + 1)$ transformed).

Model classification accuracy was also evaluated using the validation dataset. According to the confusion matrix produced, this model correctly classified 80.3% of all observations. Among observed ignitions, 78.2% were correctly predicted by the model, and 82.7% of non-ignitions were also well classified. The omission and commission errors for both ignition and non-ignition were not very high. Omission error of ignition events was 16.7%, representing the percentage of observed ignitions that were not predicted by this model, and the commission error was 21.8%, representing the percentage of expected ignitions that were not observed. The comparison between the model accuracy obtained with the training and the validation datasets revealed very small differences (79.8 *v.* 80.3%). The AUC using the validation dataset was 87.2% (s.e. = ±0.001, $P < 0.001$).

In a subsequent step, we also evaluated the possibility of developing a simpler model to predict ignition occurrence, as it would allow easier use by managers. A model with fewer variables is expected to be more stable and easily generalized; however, the more variables are included in the model, the higher will the estimated standard errors and the model dependence on the observed data (Hosmer and Lemeshow 1989). Thus, a new logistic model was developed without using the land cover variable, which is more likely to change through time than the other variables.

The performance of the model using only three variables (Table 2) was not much different from the model using four

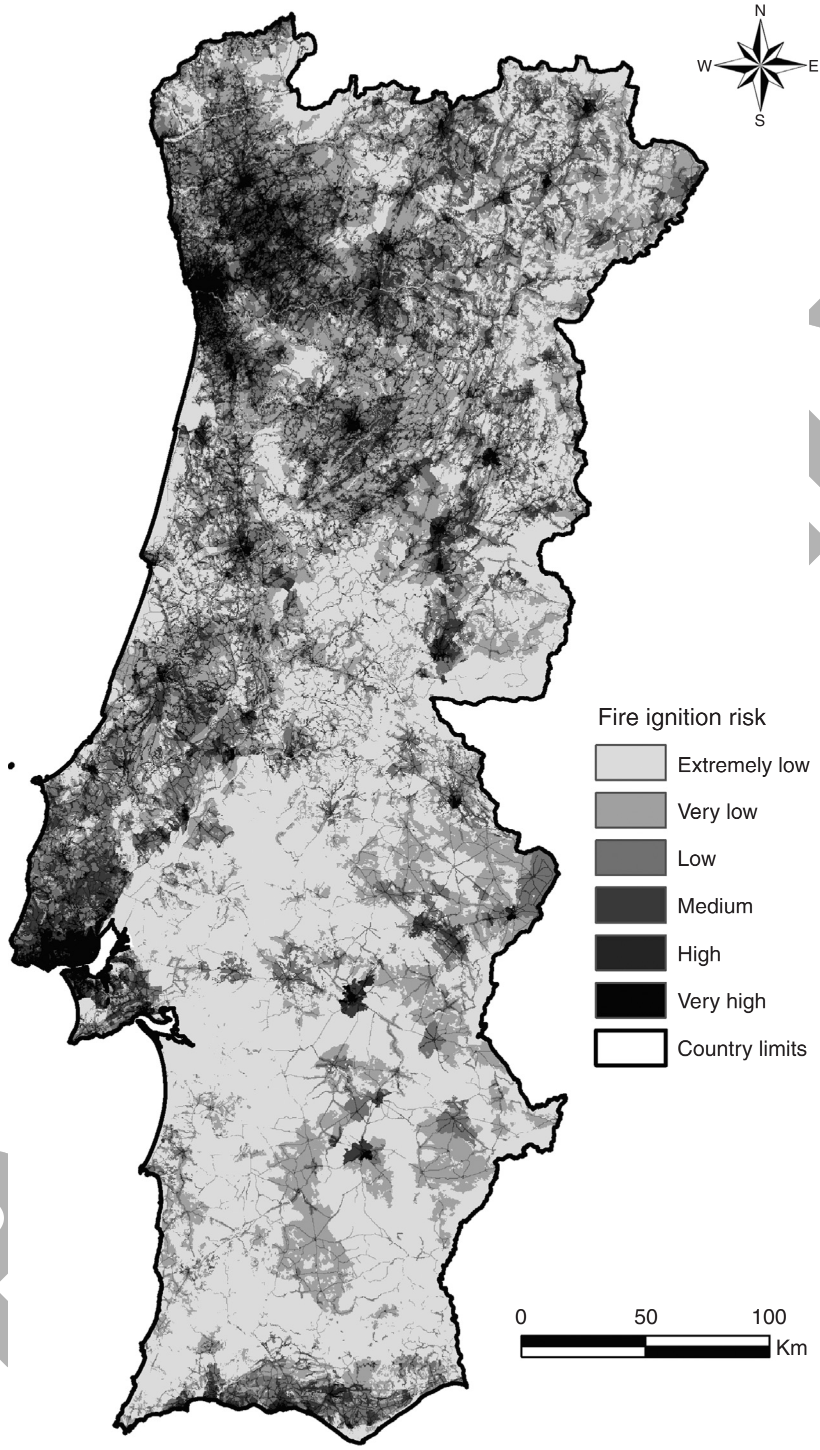

Fig. 3. Fire ignition risk map produced for the entire Portuguese mainland**[AQ4]**.

variables. The simplest model presented a global accuracy of 77.1% with both the training and the validation datasets. The omission error of ignition events was 23.2% and the commission error was 22.7%. The AUC was 0.847 (s.e. $= \pm 0.001$, $P < 0.001$) with the training dataset, and 0.849 (s.e. $= \pm 0.001$, $P < 0.001$) with the validation dataset.

The new model obtained is represented by the following equation:

$$P_2 = 1/(1 + e^{-(-5.890+0.875 Pop\_D - 0.214 D\_roads + 0.473 Elev)}).$$

In order to evaluate model stability during the study period, we compared the coefficient values and the model performance

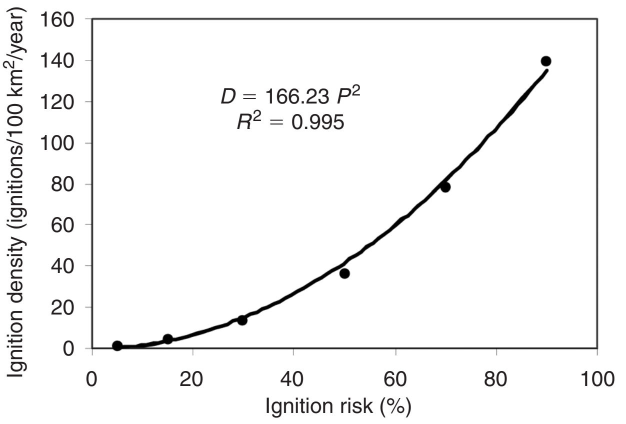

Fig. 4. Density of ignitions (D) as an average for 2001–05, as a function of P_1 (P, probability of ignition occurrence), as computed in the *Fire ignition models* section**[AQ5]**.

of the global model (fitted with data from the 5-year period) with those from separate models for each of the 5 years, and verified that both coefficients and performance (AUC range = 0.832–0.855) were very similar. The same constancy was observed in the model using four variables.

Fire ignition risk map

The fire ignition model obtained using four explanatory variables was spatialized in a GIS environment. The map obtained representing the relative probability that an ignition occurs at a given location (Fig. 3) was classified into six risk classes ranging from extremely low to very high. The representation of each class in the country was the following: extremely low (covering 42.1% of the territory), very low (23.0%), low (18.9%), medium (8.8%), high (4.6%) and very high (2.6%). The areas with high and very high risk of ignition occurrence, representing only 7.2% of the country, are mainly located in the north and central coastal regions, where human presence and activity is more important.

We evaluated the performance of the ignition risk map using a confusion matrix, as we did previously for model evaluation. The map produced also showed good predictive results, with a global accuracy of 79.1%. With the objective of describing the map prediction ability in a more detailed way, we overlaid the validation ignition points with the map, recording the exact risk class where the ignitions were located, verifying that more than 50% of them were located in high and very high risk classes although they only represent 7% of the country area (i.e. ignitions occurred at these classes with a much higher frequency than would be expected in a random distribution). However, only 10% of the ignitions were located in the two lowest risk classes, which cover more than 65% of the country area.

Average ignition density in each risk class (ignitions per 100 km^2 per year) was also computed to give a better idea of the potential number of ignitions that managers can expect. In a 1-year period, the ignition density is expected to be ∼139 times higher in the areas with very high risk than in areas with extremely low risk (covering 42% of the country), and four times higher than in areas with medium ignition risk. There is a strong relationship between the obtained fire ignition risk (P) and the real average density of ignition (D) (Fig. 4). The combination of the model developed for P_1 and the relationship between P_1 and D allows a general comprehensive final equation that predicts the average density of ignitions directly from the explanatory variables as:

$$D = 166.23/(1+e^{-(-7.833+0.820Pop\_D-0.166D\_roads+0.585Elev+2.455Urb+1.672Agr+0.388For+0.439Shr+0.426Spa)})^2.$$

Discussion and conclusions

Determinants of fire ignition

Portugal is the southern European country with the highest density of fire ignitions (EC 2007). Results of the present study show that, as hypothesized, human presence and activity are the key drivers of ignitions in Portugal. The selected explanatory variables, namely population density, distance to roads, land cover and elevation, were all highly significantly related with the spatial distribution of ignitions.

Population density was the most important variable in our model. This variable showed a positive influence on ignition occurrence, meaning a higher probability of ignition in the more populated areas, as was previously hypothesized. In regions where most fires are human-caused, this is a logical result, and in several other studies, population density was found to positively related to wildfire ignitions (e.g. Cardille *et al*. 2001; Mercer and Prestemon 2005). Also according to prior expectations, distance to roads showed a negative influence on ignitions distribution, meaning a decreasing probability of ignition occurrence with increasing distance to roadways. This result is also consistent with other authors' findings (e.g. Vega-Garcia *et al*. 1996; Romero-Calcerrada *et al*. 2008). Roadways increase human access to several areas, including those with low population density, and we think that this is the main reason for the higher frequency of ignitions near roads. In our study, 98% of ignitions occurred less than 2 km from the nearest road and 85% were within a distance of 500 m.

Land cover showed a strong influence on fire ignitions, and other authors also found this variable to be important (e.g. Cardille and Ventura 2001; Vasconcelos *et al*. 2001; Yang *et al*. 2007). Most ignitions were located in agricultural and urban–rural interspersed areas (85%), and only 15% occurred in forested or uncultivated areas, although they cover 50% of the country. Our results indicate that agriculture is a very important factor influencing fire starts and are in accordance with previous investigations of fire causes in Portugal and Spain, which concluded that a large proportion of wildfires (both in terms of number and of resulting area burned) is due to agricultural activities and mainly to agricultural burns (DGF 2003; MMA 2007). Also contributing to this high ignition incidence is probably the fact that more people are usually present in these areas, and that the herbaceous vegetation in many Mediterranean agricultural areas is easier to ignite and propagate fires in than other fuel types, especially during the summer when fuel moisture is very low. The high incidence of ignitions in urban–rural areas may also be explained by the co-occurrence of agricultural activities and a higher human presence. Forests, shrublands and sparsely vegetated areas also showed a positive influence on ignition occurrence, but their influence was considerably lower. The ignition frequency of these land cover classes was approximately three times lower than would be expected in a random distribution. Although small differences exist between forests, shrublands and sparsely vegetated areas, we can rank land cover classes in terms of the relative probability of ignition occurrence in the following way: urban–rural > agriculture > shrubs > sparsely vegetated > forests > wetlands. Cardille and Ventura (2001) also found that fires in the NW USA were more likely to start in non-forest than within forests, independently of the resulting fire size, and other authors reported that an important proportion of wildfires in southern Europe start in non-forested areas (Badia-Perpinyã and Pallares-Barbera 2006; EC 2007). However, these patterns can be completely distinct in regions with different characteristics (Yang *et al*. 2007).

Elevation was found to have a positive influence on ignition distribution. This effect may be due to the fact that there are some human activities at higher altitudes, such as the renovation of pastures for livestock, which are known to cause frequent burns in the Iberian Peninsula, and to be one of the major factors contributing to the total area burned (DGF 2003; MMA 2007). For example Badia-Perpinyã and Pallares-Barbera (2006) also found a higher ignition frequency at higher elevation in a rural area of NE Spain. Additionally, and according to some authors, lightning-caused ignitions are also more likely to occur at higher elevations (Vazquez and Moreno 1998; Kilinc and Beringer 2007).

Predictive ability of the model

The model obtained using four variables showed good predictive ability when applied to the validation dataset. Analysis of the ROC curve indicates 87.2% concordance between predicted probabilities and observed outcomes, while with the confusion matrix method, we achieved a global accuracy of 80.3%.

Although this model was developed at a national level, and thus with large human and natural variability, the results were quite good when compared with other logistic regression models developed to predict fire occurrence. For example, Vasconcelos *et al*. (2001) modelled the spatial occurrence of fire ignitions in five Portuguese municipalities ($\sim$1.6% of the country area), obtaining a global accuracy of 73.9% (omission and commission errors were respectively 22.1 and 48.6%). Chuvieco *et al*. (1998, 2003) developed models to predict the occurrence of large fires in southern Europe, obtaining a global accuracy of 60.0% (omission and commission errors for observed fires were respectively 37.5 and 51.7%). In another study, Vega-Garcia *et al*. (1995) also used logistic regression to predict the daily occurrence of fires in Alberta (Canada), obtaining a global accuracy of 74.1%; these authors obtained an omission error of 25.9% for fire-days and a very high commission error (95.1%). Although only the first of the models referred to is completely comparable with the one developed in the present work, these can give an idea of the accuracy levels achieved in modeling fire occurrence.

A simpler model, using only three variables, was also developed. In this model, the land cover variable was not included, as it is more complex to compute, and because this variable is expected to suffer more changes in the short to medium term. As expected, its global accuracy was lower than in the more complex model (77.1 and 80.3%, respectively), but the differences concerning both global accuracy and omission or commission errors were not sufficiently important to discourage its use as an alternative solution.

Additionally, it was confirmed that models built for different years had a very similar performance, which seems to indicate that little spatial variability is to be expected in a 5-year period, and that models developed using information from 1 or more years are expected to achieve similar results. The final equation expressing D as a function of the explanatory variables has two different components: a geographical component expressed by P_1, and a coefficient, representing the maximum local density of ignitions, estimated at 166.23 ignitions per 100 km^2 per year for the period 2001–05, but that can vary yearly in accordance with varying weather and social behavior.

Implications for management

When using the model or the map produced, the following issues should be kept in mind, particularly in terms of fire prevention and definition of priorities for firefighting:

(1) Special attention should be given to large forested areas with high or very high risk, where ignitions can easily occur and produce large forest fires. However, areas with low or medium ignition risk can also be very susceptible to wildfires because of high biomass accumulation. Previous analysis (e.g. Rego *et al*. 2004) showed that the areas of the country with the highest density of ignitions do not coincide with those where larger fires occur. The increasing human migratory fluxes from inland to coastal areas have led to a situation of land abandonment, contributing to reducing the risk of fire ignition in many inland regions. However, this situation simultaneously increased fuel accumulation and consequently increased fire hazard. At the same time, many of these areas have a low population density and are more distant from roads, being less accessible to firefighters. Thus, these areas should be analyzed in fire planning and surveillance operations by simultaneously evaluating the fire hazard.

(2) Attention should also be given to non-forested areas with a higher ignition risk, especially when they are near potentially hazardous areas that are important to preserve (owing to their social, economical or ecological interest). Nowadays, in several southern European countries (e.g. Portugal, Spain, Greece), a large proportion of both number of fires and area burned is not restricted to forested land (Moreno *et al*. 1998; EC 2007; MMA 2007), and many fires starting in agricultural areas or shrublands can rapidly spread and propagate into forest stands or to other areas with a higher fire hazard (DGF 2003).

(3) Because of the strong positive influence of population density on fire ignition risk, it can be seen that the areas corresponding to the centre of the larger cities (namely Lisboa and Porto) have been classified with a high risk of ignition, although we know that in fact very few ignitions occur in these highly urbanized areas because they have almost no fuels to ignite . However, this kind of problem will obviously be eliminated if the fire ignition map is used in association with a fire hazard map (or alternatively, by overlaying it with a fuel load or fuel models map and reclassifying the areas without vegetation).

(4) To use the information presented at the regional level, model adjustments, including updated and more detailed data or the inclusion of other variables, can be required to improve the accuracy of predictions. Variables could include, for example, summer population density in regions where tourism highly increases the population during the fire season, as happens in Algarve (southern Portugal), or cattle density in areas where pastoralism is an important activity.

In this work, we demonstrated the feasibility of modeling and mapping ignition risk at a national level using a limited number of easily obtainable variables. We consider that both the models developed and the ignition risk map produced have enough predictive accuracy to be used in predicting the likelihood of ignition occurrence in Portugal. Nowadays, many fire-management decisions are still exclusively based on the factors influencing fire spread and suppression difficulty. However, and as resources are limited, it is important to define priorities among areas. Under similar fuel, topographic or weather conditions, the areas with higher ignition risk should be given priority for surveillance (Vasconcelos *et al*. 2001; Chuvieco *et al*. 2003). The obtained models and the ignition risk map are now available and can be easily used in a GIS environment to integrate fire danger estimation systems, helping managers to define vigilance priority areas for ground patrol units, or to define priority locations for new lookout tower installation (Catry *et al*. 2004; Rego and Catry 2006; Catry *et al*. 2007*b*). These results can also be used for more informed decisions of firefighting resources allocation, or to optimize fire prevention public campaigns, by indicating the most problematic areas.

In spite of good results obtained, the prediction model can probably be improved by including more accurate and updated information, or by including some additional explanatory variables. It would be very useful if the national authorities could improve the ignition location accuracy, especially if the objectives are to model this event at the local or municipal level. We expect that this study, the first performed in Portugal at a national level, will encourage further research needed on this topic, which could include for example analyzing the characteristics of ignitions that resulted in large fires, or measures of temporal fire activity, such as predictions of number of fires per day. Additional research on spatial and temporal factor variability of different causes should also be useful.

Acknowledgements

We acknowledge the Portuguese Forest Services (DGRF) for all collaboration and for making available the wildfire database. We acknowledge Paula Lopes, António Nunes and Vasco Nunes for their help on preliminary data processing. We also acknowledge several important comments made by three anonymous reviewers, which contributed to improve this paper. Part of this study was supported by the European Commission under the 6th Framework Program through the Integrated Project 'Fire Paradox' (contract no. FP6–018505), and by Instituto de Financiamento da Agricultura e Pescas through the project 'Recuperação de Áreas Ardidas'.

References

Amatulli G, Pérez-Cabello F, de la Riva J (2007) Mapping lightning/human-caused wildfire occurrence under ignition point location uncertainly. *Ecological Modelling* **200**, 321–333. doi:10.1016/J.ECOLMODEL.2006.08.001

Badia-Perpinyã A, Pallares-Barbera M (2006) Spatial distribution of ignitions in Mediterranean periurban and rural areas: the case of Catalonia. *International Journal of Wildland Fire* **15**, 187–196. doi:10.1071/WF04008

Bonazountas M, Kallidromitou D, Kassomenos PA, Passas N (2005) Forest fire risk analysis. *Human and Ecological Risk Assessment* **11**, 617–626. doi:10.1080/10807030590949717

Cardille JA, Ventura SJ (2001) Occurrence of wildfire in the northern Great Lakes Region: effects of land cover and land ownership assessed at multiple scales. *International Journal of Wildland Fire* **10**(2), 145–154. doi:10.1071/WF01010

Cardille JA, Ventura SJ, Turner MG (2001) Environmental and social factors influencing wildfires in the Upper Midwest, USA. *Ecological Applications* **11**(1), 111–127. doi:10.1890/1051-0761(2001)011[0111:EASFIW]2.0.CO;2

Carreiras JMB, Pereira JMC (2006) An inductive fire risk map for Portugal. In 'Proceedings of the 5th International Conference on Forest Fire Research', 27–30 November 2006, Figueira da Foz, Portugal. (Ed. DX Viegas) (CD ROM) (Associação para o Desenvolvimento da Aerodinâmica Industrial (ADAI): Coimbra, Portugal)

Catry FX, Almeida RM, Rego FC (2004) Produção de cartografia de visibilidades para Portugal continental. A importância da sua utilização na vigilância contra incêndios florestais. *Silva Lusitana* **12**(2), 227–241. [In Portuguese]

Catry FX, Damasceno P, Silva JS, Galante M, Moreira F (2007*a*) Spatial distribution patterns of wildfire ignitions in Portugal. In 'Proceedings of the 4th International Wildland Fire Conference**[AQ1]**', 13–17 May 2007, Seville, Spain. (CD ROM)

Catry FX, Rego FC, Santos T, Almeida J, Relvas P (2007*b*) Forest fires prevention in Portugal – using GIS to help improving early fire detection effectiveness. In 'Proceedings of the 4th International Wildland Fire Conference', 13–17 May 2007, Seville, Spain. (CD ROM)

Chou YH (1992) Management of wildfires with a geographical information system. *International Journal of Geographical Information Systems* **6**, 123–140. doi:10.1080/02693799208901900

Chuvieco E, Salas J, Barredo JI, Carvacho L, Karteris M, Koutsias N (1998) Global patterns of large fire occurrence in the European Mediterranean Basin: a GIS analysis. In 'Proceedings of the 3rd International Conference on Forest Fire Research – 14th Conference on Forest Fire Meteorology. Vol II'. (Ed. DX Viegas) pp. 2447–2462. (Associação para o Desenvolvimento da Aerodinâmica Industrial (ADAI): Coimbra, Portugal)

Chuvieco E, Allgöwer B, Salas J (2003) Integration of physical and human factors in fire danger assessment. In 'Wildland Fire Danger Estimation and Mapping. The Role of Remote Sensing Data'. (Ed. E Chuvieco) Vol. 4, pp. 197–218. (World Scientific Publishing: Singapore)

Congalton R (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* **37**, 35–46. doi:10.1016/0034-4257(91)90048-B

DGF (2001) 'Inventário florestal nacional. Portugal continental. 3a Revisão, 1995–1998.' (Direcção-Geral das Florestas: Lisbon) [In Portuguese]

DGF (2003) Determinação das causas dos incêndios florestais em 2002. Direcção-Geral das Florestas Report. (Lisbon, Portugal) [In Portuguese]

DGRF (2006) Incêndios florestais – Relatório de 2005. Divisão de Defesa da Floresta Contra Incêndios. Direcção-Geral dos Recursos Florestais Report. (Lisbon, Portugal) [In Portuguese]

EC (2007) Forest fires in Europe 2006. Official Publication of the European Commission, EUR 22931 EN. (Ispra, Italy)

ESRI (2005) '*ArcGIS 9.1*. Software.' (Environmental Systems Research Institute: Redlands, CA)

FAO (1986) Wildland fire management terminology. Food and Agriculture Organization of the United Nations, FAO Forestry Paper 70. (Rome, Italy)

Finney MA (2005) The challenge of quantitative risk analysis for wildland fire. *Forest Ecology and Management* **211**, 97–108. doi:10.1016/J.FORECO.2005.02.010

Genton MG, Butry DT, Gumpertz ML, Prestemon JP (2006) Spatio-temporal analysis of wildfire ignitions in the St Johns River Water Management District, Florida. *International Journal of Wildland Fire* **15**, 87–97. doi:10.1071/WF04034

Hosmer DW, Lemeshow S (1989) 'Applied Logistic Regression.' (John Wiley and Sons: New York)

IA (2003) Atlas do Ambiente Digital. Instituto do Ambiente. Available at www.iambiente.pt/atlas/est/index.jsp [Verified 2 March 2007]

IA (2005) CORINE Land Cover 2000 Portugal. Instituto do Ambiente. (Lisbon, Portugal)

IGEOE (2005) Carta Militar Itinerária de Portugal. Instituto Geográfico do Exército. Available at www.igeoe.pt [Verified 2 March 2007]

IGP (2004*a*) Cartografia de risco de incêndio florestal. Relatório do distrito de Viseu. Centro para a Exploração e Gestão de Informação e Situações de Emergência, IGP Technical Report. (Lisbon, Portugal) [In Portuguese]

IGP (2004*b*) Carta administrativa oficial de Portugal. Instituto Geográfico Português. Available at www.igeo.pt/igeo/portugues/Frameset-egeo.htm [Verified 2 March 2007]

INE (2003) Dados estatísticos da população em Portugal – Censos 2001. Instituto Nacional de Estatística. (Lisbon, Portugal) [Database in Portuguese]

Johnson EA, Miyanishi K (2001) 'Forest Fires: Behavior and Ecological Effects.' (Academic Press: San Diego, CA)

Koutsias N, Kalabokidis KD, Allgöwer B (2004) Fire occurrence patterns at landscape level: beyond positional accuracy of ignition points with kernel density estimation methods. *Natural Resource Modeling* **17**(4), 359–375.

Kilinc M, Beringer J (2007) The spatial and temporal distribution of lightning strikes and their relationship. *Journal of Climate* **20**(7), 1161–1173. doi:10.1175/JCLI4039.1

Legendre P, Legendre L (1998) 'Numerical Ecology.' 2nd English edn (Elsevier Science BV: Amsterdam)

Loboda TV, Csiszar IA (2007) Assessing the risk of ignition in the Russian Far East within a modelling framework of fire threat. *Ecological Applications* **17**(3), 791–805. doi:10.1890/05-1476

Manel S, Williams H, Ormerod SJ (2001) Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**, 921–931. doi:10.1046/J.1365-2664.2001.00647.X

Mercer DE, Prestemon JP (2005) Comparing production function models for wildfire risk analysis in the wildland–urban interface. *Forest Policy and Economics* **7**, 782–795. doi:10.1016/J.FORPOL.2005.03.003

MMA (2007) Los incendios forestales en España. Decenio 1996–2005. Area de Defensa Contra Incendios Forestales, Ministerio de Medio Ambiente. (Madrid, Spain)

Moreno JM, Vázquez A, Vélez R (1998) Recent history of forest fires in Spain. In 'Large Forest Fires'. (Ed. JM Moreno) pp. 159–186. (Backhuys Publishers: Leiden, the Netherlands)

NASA, NIMA, DLR, ASI (2004) Digital elevation model. National Aeronautics and Space Administration; National Imagery and Mapping Agency; German space agency (DLR); Italian space agency (ASI). Available at ftp://edcsgs9.cr.usgs.gov/pub/data/srtm [Verified 8 August 2006]

NWCG (2006) Glossary of wildland fire terminology. National Wildfire Coordinating Group, PMS 205 (Boise, ID)

Norusis MJ (2002) '*SPSS* for Windows – Base System User's Guide.' (SPSS Inc.: Chicago, IL)

Nunes A, Duarte J (2006) Assessment of forest fire risk in the Serra da Estrela Natural Park (Portugal): methodological application and validation. In 'Proceedings of the 5th International Conference on Forest Fire Research', 27–30 November 2006, Figueira da Foz, Portugal. (Ed. DX Viegas) (CD ROM) (Associação para o Desenvolvimento da Aerodinâmica Industrial (ADAI): Coimbra, Portugal)

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* **133**, 225–245. doi:10.1016/S0304-3800(00)00322-7

Preisler HK, Brillinger DR, Burgan RE, Benoir JW (2004) Probability-based models for estimation of wildfire risk. *International Journal of Wildland Fire* **13**, 133–142. doi:10.1071/WF02061

Rego FC, Catry FX (2006) Modelling the effects of distance on the probability of fire detection from lookouts. *International Journal of Wildland Fire* **15**, 197–202. doi:10.1071/WF04016

Rego FC, Catry FX, Maia MJ, Santos TA, Gravato A, Castro IC, Moreira FO, Pinto PR, Almeida J (2004) Análise da rede nacional de postos de vigia em Portugal. Centro de Ecologia Aplicada Prof. Baeta Neves – Instituto de Novas Tecnologias, Technical Report (Lisbon, Portugal) [In Portuguese]

Robin JG, Carrega P, Fox D (2006) Modelling fire ignition in the Alpes-Maritimes Department, France, a comparison. In 'Proceedings of the 5th International Conference on Forest Fire Research', 27–30 November 2006, Figueira da Foz, Portugal. (Ed. DX Viegas) (CD ROM) (Associação para o Desenvolvimento da Aerodinâmica Industrial (ADAI): Coimbra, Portugal)

Roloff GJ, Mealey SP, Clay C, Barry J, Yanish C, Neuenschwander L (2005) A process for modelling short- and long-term risk in the southern Oregon Cascades. *Forest Ecology and Management* **211**, 166–190. doi:10.1016/J.FORECO.2005.02.006

Romero-Calcerrada R, Novillo CJ, Millington JDA, Gomez-Jimenez I (2008) GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). *Landscape Ecology* **23**, 341–354. doi:10.1007/S10980-008-9190-2

Rorig ML, Ferguson SA (1999) Characteristics of lightning and wildland fire ignition in the Pacific North-west. *Journal of Applied Meteorology* **38**, 1565–1575. doi:10.1175/1520-0450(1999)038<1565:COLAWF>2.0.CO;2

Sokal RR, Rohlg FJ (1987) 'Biometry.' (WH Freeman and Company: New York)

SPSS (2006) '*SPSS* for Windows.' (SPSS Inc.: Chicago, IL)

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* **240**, 1285–1293. doi:10.1126/SCIENCE.3287615

Vasconcelos MJP, Silva S, Tomé M, Alvim M, Pereira JMC (2001) Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. *Photogrammetric Engineering and Remote Sensing* **67**(1), 73–81.

Vasilakos C, Kalabokidis K, Hatzopoulos J, Kallos G, Matsinos Y (2007) Integrating new methods and tools in fire danger rating. *International Journal of Wildland Fire* **16**(3), 306–316. doi:10.1071/WF05091

Vazquez A, Moreno JM (1998) Patterns of lightning- and human-caused fires in peninsular Spain. *International Journal of Wildland Fire* **8**(2), 103–115. doi:10.1071/WF9980103

Vega-Garcia C, Woodard PM, Titus SJ, Adamowicz WL, Lee BS (1995) A logit model for predicting the daily occurrence of human-caused forest fires. *International Journal of Wildland Fire* **5**(2), 101–111. doi:10.1071/WF9950101

Vega-Garcia C, Lee BS, Woodard PM, Titus SJ (1996) Applying neural network technology to human-caused wildfire occurrence prediction. *AI Applications* **10**(3), 9–18.

Yang J, Healy HS, Shifley SR, Gustafson EJ (2007) Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. *Forest Science* **53**(1), 1–15.

Manuscript received 25 August 2007, accepted 30 January 2009

AUTHOR QUERIES

AQ1 Au: please provide publisher for this and Catry *et al*. 2007*b*
AQ2 Instruction for typesetters: kilometers needs changing to km
AQ3 Instruction for typesetters: units need the forward slash removed and the space before > removed
AQ4 Instruction for typesetters: Km should be km
AQ5 Instruction for typesetters: forward slash need removing and the units need correct spacing
AQ6 Au: is this a reference citation? If so please provide a reference. Otherwise some further explanation would be helpful.