

GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain)

Raul Romero-Calcerrada · C. J. Novillo ·
J. D. A. Millington · I. Gomez-Jimenez

Received: 6 August 2007 / Accepted: 7 January 2008 / Published online: 30 January 2008
© Springer Science+Business Media B.V. 2008

Abstract The majority of wildfires in Spain are caused by human activities. However, much wildfire research has focused on the biological and physical aspects of wildfire, with comparatively less attention given to the importance of socio-economic factors. With recent changes in human activity and settlement patterns in many parts of Spain, potentially contributing to the increases in wildfire occurrence recently observed, the need to consider human activity in models of wildfire risk for this region are apparent. Here we use a method from Bayesian statistics, the weights of evidence (WofE) model, to examine the causal factors of wildfires in the south west of the Madrid region for two differently defined wildfire seasons. We also produce predictive maps of wildfire risk. Our results show that spatial patterns of wildfire ignition are strongly associated with human access to the natural landscape, with proximity to urban areas and roads found to be the most important causal factors. We suggest these characteristics and recent

socio-economic trends in Spain may be producing landscapes and wildfire ignition risk characteristics that are increasingly similar to Mediterranean regions with historically stronger economies, such as California, where the urban-wildland interface is large and recreation in forested areas is high. We also find that the WofE model is useful for estimating future wildfire risk. We suggest the methods presented here will be useful to optimize time, human resources and fire management funds in areas where urbanization is increasing the urban-forest interface and where human activity is an important cause of wildfire ignition.

Keywords Human-caused fires · Ignition risk · Socio-economic factors · Spatial pattern · Fire risk · Weights of evidence · GIS · Wildland fire

Introduction

Much wildfire research has focused on the biological and physical aspects of fire, with comparatively less attention given to the importance of socio-economic variables. However, previous work in the Iberian Peninsula has demonstrated that the proportion of fires started by direct or indirect human causes surpass the natural causes (Moreno et al. 1998; Vázquez and Moreno 1995; WWF/Adena 2004, 2005). In Spain, where more than 95% of wildfires are caused by human activity for example (MMA 2007), estimation of ignition risk must consider the influence of human activity. The high incidence of

R. Romero-Calcerrada (✉) · C. J. Novillo ·
I. Gomez-Jimenez
School of Experimental Science and Technology,
Rey Juan Carlos University, C/ Tulipán s/n,
28933 Mostoles, Madrid, Spain
e-mail: raul.romero.calcerrada@urjc.es

J. D. A. Millington
Center for Systems Integration & Sustainability, Michigan
State University, 1405 S. Harrison Road, Suite 115 Manly
Miles Bldg, East Lansing, MI 48823-5243, USA

wildfires in Spain in recent years underlines and emphasises the importance of understanding the causes and spatial distribution of this phenomena. Increases in wildfire frequency and magnitude observed recently (MMA 2007) may be due, in part, to the increase of human populations in forested areas. There are two reasons for this population increase: (i) urban growth is encroaching upon forested areas resulting in more people living at the urban-forest interface; and (ii) recreational activity is increasing in forested areas.

Despite these findings, human factors have been given scant regard compared to physical factors (elevation, slope, temperature, rainfall, vegetation, etc.) in quantitative analyses of risk (Chou et al. 1990, 1993; de Vasconcelos et al. 2001; Yang et al. 2007). This disregard for the primary cause of wildfires in the Iberian Peninsula is owed, without doubt, to the difficulties of spatially evaluating and modelling the human component of both fire ignition and spread. However, understanding the spatial influence of human activities on the distribution of ignitions is central to managing and mitigating the ignition risk. If a fire is to be extinguished it is important to attack it early, whilst it is small, making it vital to position fire fighting resources near areas that ignite most frequently.

The seminal work in this area of research by Chuvieco and Congalton (1989) incorporated human activity into the assessment of wildfire risk by considering the proximity of zones of greater human activity near road networks and recreational areas. In recent years wildfire risk models that consider other human variables have become increasingly common (Badia-Perpinyà and Pallares-Barbera 2006; Cardille et al. 2001; Chou et al. 1993; de Vasconcelos et al. 2001; Dickson et al. 2006; Yang et al. 2007). These variables have included Distance to Recreational Areas/campgrounds, location of urban areas, population density, unemployment levels and land tenure. However, there is still much work to be done to adequately estimate the risk, and predict the occurrence, of wildfires in regions where the majority of events are ignited by human activity. We suggest that these studies have only considered a small fraction of the large number of possible socio-economic factors linked with the occurrence of wildfires in the Iberian Peninsula.

We propose a novel approach to analyze fire ignition patterns for an area in the SW of Madrid (Central

Spain). This paper focuses on ignition risk over the period 2000–2003. In an effort to better understand the conditions conducive to wildfire ignition two different spatial models are used for both four- and two-month fire seasons. Each model considers the relationship between socio-economic variables and wildfire ignition locations. Weights of evidence (WofE) GIS (Kemp et al. 2001) is used to gain understanding regarding the main causes of ignition and to build a predictive model of wildfire ignition risk.

Methods

Study area

The study area is composed of three forestry administration areas (Parque Regional del Guadarrama, San Martín de Valdeiglesias and Robledo de Chavela) situated to the southwest of Madrid (Central Spain—Fig. 1). It comprises an area of some 1808 km² (22.5%) of the Madrid Region, encompassing 45 municipalities (25% of Madrid Region).

As with many landscapes in the Mediterranean Basin, the study area is of both high social and cultural value and supports multiple land uses, including agriculture, tourism and recreational activities, and environmental services. The mountainous area along the western border has a low population density and has experienced aging in its farming population and a consequent abandonment of traditional agrarian land uses. The abandonment of crops (e.g. vineyard) and pastures has resulted in ‘old-field’ succession in these areas (Romero-Calcerrada and Perry 2004). In the east of the study area, nearer the city of Madrid, population density is greater and the main industries are secondary and tertiary in nature. Urban development and a shift from pastureland to scrubland, as a result of land abandonment and arable land lying fallow, was the dominant change in the study area during the period 1984–2006. Urban growth and recreational land uses have also increased over recent years due to the aesthetic value of the landscape, relative low cost of housing and proximity to Madrid. A probable consequence of land cover and land use changes is an altered fire risk (Millington 2005; Romero-Calcerrada and Perry 2002; 2004).

The study area is also interesting from the point of view of wildfire. Analysis of the wildfire database

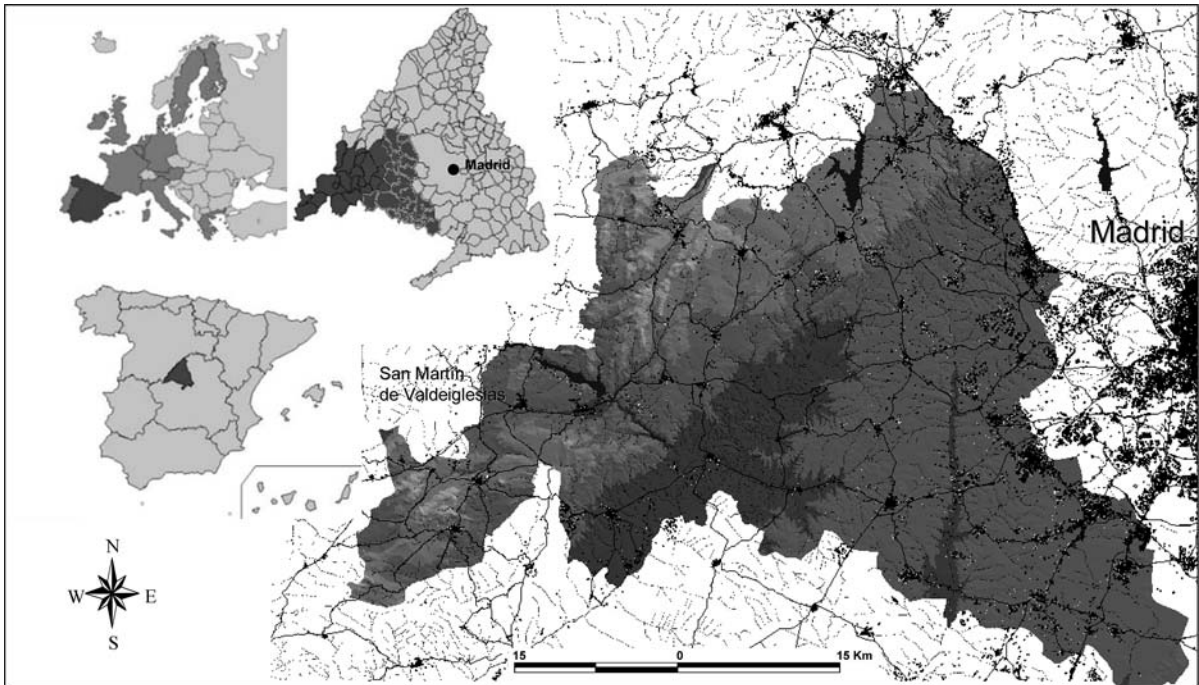


Fig. 1 Study area location: SW of Madrid (Central Spain)

used in this study showed that the majority of wildfire ignitions are caused by human activity (95% of events between 1996 and 2005). Furthermore, the number of fires increased but the total burned area decreased in this period, possibly reflecting increased fire-fighting efforts in these areas.

Data sources

In this study, we used 11 independent variables (Table 1), that included five socio-economic variables and six spatial variables (spatial relationships of variables). The spatial variables were defined to represent human access across the study area and the spatial pattern of human land use. National and regional statistics were examined to find data on the main socio-economic aspects that could be used to characterize ignition risk. All variables were selected because of their influence on wildfire ignition risk. Maps of independent variables were categorized using the Natural Breaks classification and Buffer Distances function in ArcViewTM 3.2. All maps were included in ArcViewTM 3.2, at 1:50,000 scales in UTM projection, and rasterized with a 25 m resolution.

We used digital reports of fire ignitions occurring between 2000 and 2005 for the SW of Madrid. We obtained the database directly from the Sección de Defensa Contra Incendios Forestales of the Regional Government of Madrid. In these records, ignition points are referenced to UTM coordinates and provide information about the timing (day/month/year) of events and burned area (tree forest areas, pastures/bush area, non-forest areas). Ignition point data were subset (Table 2) to the years 2000–2003 for a four- (Model 1) and two-month fire season (Model 2). The remaining 2 years in the dataset (2004–2005) were used as testing data. Model 1 months were June, July, August and September (“official” fire season). Model 2 considers July and August alone as human activity is greatest in the countryside, with approximately 55% of all wildfires occurring in this period.

Weights of evidence

Weights of evidence (Kemp et al. 2001) is a data-driven, discrete multivariate statistical method that uses Bayesian principles for integrating multiple evidence (predictor or independent) variables and

Table 1 Description of socio-economic and spatial variables

Abbreviation	Group	Description (Units)	Sources
Dy_Pop	Socio-economic data	Density of Population (Inhab/ha of Urban Area)	INE (2001). Owner Elaboration
Dy_2Housing	Socio-economic data	Density of Secondary Housing (Secondary Housing/ha of Urban Area)	INE (2001) Owner Elaboration
Dy_Cattle	Socio-economic data	Density of Cattle (Cattle/ha of pastures)	INE (1999) Owner Elaboration
Dy_Sheep	Socio-economic data	Density of Sheep (Sheep/ha of pastures)	INE (1999) Owner Elaboration
Dy_Goats	Socio-economic data	Density of Goats (Goats/ha of pastures)	INE (1999) Owner Elaboration
D_Urb	Spatial relation	Distance to Urban Areas (m)	SCR (2000)
D_Ind	Spatial relation	Distance to Industrial Areas (m)	SCR (2000)
D_Roads	Spatial relation	Distance to Roads (Highways etc.) (m)	SCR (2000)
D_Tracks	Spatial relation	Distance to Tracks: rural path etc. (m)	SCR (2000)
D_Camping	Spatial relation	Distance to Camping (m)	D. G. del Medio Natural (1997)
D_Recreat	Spatial relation	Distance to Recreational Areas (m)	D. G. del Medio Natural (2000)

Table 2 Number of Ignition Points 2000–2005 in the study area

Seasons	Subset	Models	Number ignitions
Jun/Jul/Aug/Sep (2000–2003)	Known occurrences	Model 1	248
Jul/Aug (2000–2003)	Known occurrences	Model 2	169
Jun/Jul/Aug/Sep (2004)	Testing data	Holdout sample to validate Model 1	93
Jul/Aug (2004)	Testing data	Holdout sample to validate Model 2	51
Jun/Jul/Aug/Sep (2005)	Testing data	Holdout sample to validate Model 1	75
Jul/Aug (2005)	Testing data	Holdout sample to validate Model 2	50
Jun/Jul/Aug/Sep (2004–2005)	Testing data	Holdout sample to validate Model 1	168
Jul/Aug (2004–2005)	Testing data	Holdout sample to validate Model 2	101

conditional probabilities to determine the relative importance of these variables on a dependent variable. Particularly, WofE provides a measure of the spatial association between maps of independent variables and dependent variable point data by using Bayes' probability theorem (Bonham-Carter 1994; Bonham-Carter and Agterberg 1999; Bonham-Carter et al. 1989). In the case of the potential distribution of wildfire ignition, a series of evidence maps consisting of a set of spatial human datasets (i.e. categorical maps of Distance of Roads, Urban Areas, etc.) maps are created for ignition occurrences.

One of the most important concepts used in WofE is the idea of prior and posterior probability. The prior probability ($P\{D\}$) is the probability of occurrence of dependent data variable D studied without consideration of any known evidence information. The posterior probability can be calculated from the prior probability $P\{D\}$ by:

$$P\{D|B_i\} = \frac{P\{B_i \cap D\}}{P\{B_i\}} = P\{D\} \frac{P\{B_i|D\}}{P\{B_i\}} \quad (1)$$

$$P\{D|\bar{B}_i\} = \frac{P\{\bar{B}_i \cap D\}}{P\{\bar{B}_i\}} = P\{D\} \frac{P\{\bar{B}_i|D\}}{P\{\bar{B}_i\}} \quad (2)$$

where $P\{D|B_i\}$ is the posterior probability of an occurrence given the class i of a predictor theme B ; and $P\{D|\bar{B}_i\}$ is the posterior probability of an occurrence given the absence of class i in theme B . The independent variable maps are combined using Bayes' rule in a multi-map overlay operation, where the prior probability ($P\{D\}$) of an occurrence is updated by the addition of independent variables and their weights to produce a single posterior probability, $P\{D|B_i\}$, map of occurrence. The posterior probability map produced is a map of potential ignition distribution that reflects the spatial distribution of previously observed events

(Aspinall 1992; Aspinall and Veitch 1993; Guisan and Zimmermann 2000; Milne et al. 1989; Skidmore 1989; Tucker et al. 1997). For a more detailed description of the mathematical basis of the Bayesian WofE model, see Agterberg et al. (1990), Bonham-Carter (1994) and Bonham-Carter and Agterberg (1999).

The WofE model, implemented previously in ArcView™ 3.2 by Kemp et al. (2001), has four main steps:

1. Calculate weights for each independent variable (evidence map)
2. Generalize the evidence map
3. Applying a conditional independence test
4. Create a predictive map ignition risk

Calculating weights for each independent variable

Observed ignition locations (occurrence points) are used to calculate the weights for each independent variable, one weight per class, using the overlap relationships between the points and the various categories in the evidence maps. Socio-economic data weights were calculated using a ‘categorical weights’ approach, with categorical weights applied to each class. In proximity analysis (e.g. distance to roads), an ‘ascending cumulative’ approach was used to weight spatial distance classes (cumulative from low distances to high). These methods were chosen based on information available from the observed location(s) of ignition points and the evidence themes.

The WofE for the class i of predictor variable B are defined as:

$$W_i^+ = \log_e \frac{P\{B_i|D\}}{P\{B_i|\bar{D}\}} \quad (3)$$

and

$$W_i^- = \log_e \frac{P\{\bar{B}_i|D\}}{P\{\bar{B}_i|\bar{D}\}} \quad (4)$$

If the spatial association is greater than would be expected at random, W^+ is positive and W^- is negative. W^+ and W^- give a unit-less measure of spatial association between a set of occurrence points and an evidence class.

The Contrast (C) for the class i between these measures is given by:

$$C_i = W_i^+ - W_i^- \quad (5)$$

A larger C value indicates stronger spatial association between the evidence and dependent data map. The C can be used to divide the data into different classes of spatial association. A high C value is produced by strong spatial association (i.e. a large number of occurrences in a given category of an evidence map). However, the uncertainty in weights increases with a diminishing number of data points and in some cases the C value can become meaningless (Carranza and Hale 2000, 2002). The Studentized value of C (C_S) is a useful measure in this instance (Bonham-Carter 1994). Calculated as the ratio of C to the standard deviation of C , C_S serves as an informal test that C is significantly different from zero, or that the contrast is ‘real’ (Bonham-Carter 1994; Carranza and Hale 2000, 2002). The C_S is also helpful and more useful than C for choosing the cut-off when categorising independent variables into evidence maps (Carranza and Hale 2000, 2002) because it shows the contrast relative to the certainty or uncertainty due to the Contrast (Bonham-Carter 1994).

Generalizing the evidential theme

WofE uses generalize weights C and C_S to obtain binary or a few multi-class evidence themes, as reduced numbers of independent variable classes results in more robust estimations (Kemp et al. 1999). Once weights, C , and C_S have been calculated for each class of each categorical evidence map, the appropriate breaks to generalize the data must be decided upon. In this study we use multi-class evidence data, as the use of a few classes in WofE allows more precise modelling, using values of C_S as the cut-off point for the grouping process. As it was unclear whether we had ‘small’ or ‘large’ areas of occurrence, we adopted a conservative attitude using C_S . For a positive spatial association, C_S will have a positive value. Positive spatial association produces high positive values of C_S which in turn means high predictive power. Values close to 0 mean lower predictive power. We used the following criteria for grouping: Group **W0** for $C_S < 1.96$; Group **W1** for

$1.96 \leq C_S < 3$; Group **W2** for $3 \leq C_S < 4$; Group **W3** for $4 \leq C_S < 5$; and Group **W4** for $C_S \geq 5$. Thus, we have used C_S to aggregate classes in different predictive groups (W0 to W5). This generalization rule is repeated for each theme used as evidence to obtain the predictive maps. Also, C_S is also used for testing the hypothesis that C is significantly different to zero.

Applying of conditional independence test

The predictive maps assume the conditional independence of the evidence maps. Violation of this assumption can result in an over- or under-estimation of the weights. Because of this, the Pairwise Conditional Independence Test and Overall Test of Conditional Independence (described in Bonham-Carter 1994) were applied.

Upon examining pairwise conditional independence tables, conditional dependence was found between combinations of evidence maps. Therefore it was necessary to select only some of the evidence maps to produce the predictive models. Our criterion was to select the combination of evidence maps with greatest C_S and to include the greatest number of evidence maps possible. As an alternative to discarding some independent variables as we have here, Dickson et al. (2006) attempted the combination of dependent evidence maps. However, this method requires further research and was not pursued here.

Creating a predictive map of human-based ignition risk

A *posterior probabilities* map is generated by combining the conditionally independent evidence maps. The resulting predictive map is obtained as results of the ratio of *posterior probabilities* ($P\{D|B\}$) and *prior probabilities* ($P\{D\}$). In our case, this means the predicted to expected wildfire ignition ratio. High uncertainty areas are masked out because “Studentized” posterior probability ($P\{D|B\}/\sigma_{\text{Total}}$) < 1.5 indicates too much uncertainty (Bonham-Carter et al. 1989). The predictive map was classified into three categories based on Carranza and Hale (2000):

1. High predictive—when the ratio ($P\{D|B\}:P\{D\}$) > 2 and ($P\{D|B\}/\sigma_{\text{Total}}$) > 1.5 .

2. Medium predictive—when $1 < (P\{D|B\}:P\{D\}) < 2$ and ($P\{D|B\}/\sigma_{\text{Total}}$) > 1.5 .

3. Low predictive—when ($P\{D|B\}:P\{D\}$) < 1 and ($P\{D|B\}/\sigma_{\text{Total}}$) > 1.5 .

Models validation

Model evaluation is an essential step when modelling and, ideally, models should be tested with data independent of that used to develop the model (Fielding and Bell 1997; Guisan and Zimmermann 2000). Here we use a three holdout sample approach for each model (Hooten et al. 2003) to assess the performance of the models for predicting ignition risk (Table 2). Model 1 and Model 2 were created from the 2000–2003 dataset of wildfire ignitions. The 2004, 2005 and 2004–2005 observed ignition point data were used to evaluate (validate) these predictive models. The three validation data set were used to estimate the robustness of models against the inter-annual variation in ignition occurrence.

We have developed a presence-only model to avoid the problems of pseudo-absences; in the case of modelling wildfire ignitions a presence only assessment is a more realistic approach (Boyce et al. 2002; Elith et al. 2006; Hirzel et al. 2006; Pearce and Boyce 2006). We applied the area-adjusted frequencies (AAF = Freq. Predicted/Freq. Expected ratio) to assess our models. For further details, see Hirzel et al. (2006), Pearce and Boyce (2006) and Boyce et al. (2002). A useful model is one for which the predicted probability of occurrence is higher than the Expected probability of occurrence.

Results

Weights, contrast and studentized C for each evidence map

We tested our hypotheses by quantifying the spatial associations between socio-economic data maps and ignition points. In general, $W^+ > 2$ are extremely predictive; $1 < W^+ \leq 2$ are strongly predictive; $0.5 < W^+ \leq 1$ are moderately predictive and $0 < W^+ \leq 0.5$ are mildly predictive (Kemp et al. 1999). These statistical values are useful to analyze

and understand the characteristics of the ignition points in each period within the study region.

The spatial association of Density of Population (Table 3) with presence of ignition is mildly and moderately predictive. These classes are statistically significant ($\alpha = 0.05$). The majority of the classes have negative high values of W^+ , only the class 14–23 present positive values. For Density of Secondary Housing, the results suggest that the presence of Secondary Housing has a direct and important relationship with the location of wildfire ignition. In relation to livestock, we find that the moderate densities of cattle (0.14–0.39) and null densities of sheep (<0.002) and goats (0) have positive high values of W^+ . This means that the

presence of livestock has an inverse relationship with the ignition points in our study area. The Distance-buffer from Recreational Areas analysis shows that wildfire ignition occurs in areas closer to recreational areas. Only the <4500 m class of Model 1 shows statistical significance ($\alpha = 0.05$), but the value of W^+ suggests low predictability.

Distance of Urban Areas shows insignificant values only for distances up to 50 m (Fig. 2a). The W^+ is moderately predictive in the cumulative classes 0–150 and 0–250. This trend is reversed for distances >350 m. These results suggest that the ignition points have a strong relationship with the proximity to urban areas. For Distance to Industrial Areas, the results show that the wildfire ignition occurs in areas

Table 3 Summary of W^+ for evidence themes in the Model 2000–2003

Dy_Pop	Classes	< 5.84	5.84–8.7	8.7–12.3	12.3–14.37	14.37–23.04	23.04–38.76	38.76–60.8	>60.8
(Inhab / ha of Urban area)	Model 1	-0.45	-0.04	0.13	-0.56	0.43	-0.25	0.21	-0.69
	Model 2	-0.58	0.11	-0.08	-0.52	0.47	-0.31	0.23	-0.87
Dy_2Housing	Classes	< 0.87	0.87–1.97	1.97–2.94	2.94–4.24	4.24–5.59	5.59–7.56	7.56–10.12	> 10.12
(secondary housing/ ha of Urban area)	Model 1	-1.06	-0.31	-0.79	0.26	0.20	0.13	-0.28	0.84
	Model 2	-0.83	-0.31	-0.86	0.25	0.20	0.15	-0.30	0.78
Dy_Cattle	Classes	< 0.04	0.04–0.14	0.14–0.39	0.39–0.75	0.75–1.92	1.92–2.78	2.78–6.83	> 6.83
(Cattle/ ha of Pastures)	Model 1	-0.08	-0.13	0.47	-0.26	-0.89			-1.90
	Model 2	0.02	-0.27	0.53	-0.44	-1.76			-1.51
Dy_Sheep	Classes	< 0.02	0.02–0.06	0.06–0.13	0.13–0.22	0.22–0.45	0.45–1.05	1.05–1.71	> 1.71
(Sheep/ ha of Pastures)	Model 1	0.48	-0.45	-0.47	-0.15	-0.68	-0.41	-0.58	0.00
	Model 2	0.46	-0.46	-0.31	-0.11	-0.61	-0.52	0.00	0.00
Dy_Goats	Classes	0	0–0.01	0.01–0.02	0.02–0.03	0.03–0.05	0.05–0.17	0.17–0.4	>0.4
(Goats/ ha of Pastures)	Model 1	0.28	0.12	-0.28	-0.20	-1.10	-0.10	-0.55	0.00
	Model 2	0.35	-0.02	-0.22	-0.30	-1.28	0.07	-0.86	0.00
D_Recreat	Classes	0–500	0–1500	0–2500	0–3500	0–4500	All distances		
(m)	Model 1	0.80	0.31	0.06	0.12	0.15	0.00		
	Model 2	0.25	0.29	0.08	0.07	0.09	0.00		

Strongly predictive classes are marked in Dark Grey; moderately predictive are marked in light Grey. Cs serves as a guide to the significance of the spatial association. Cs values greater than 1.96 indicate that the hypothesis that $C = 0$ can be rejected at $\alpha = 0.05$ (these values are highlighted in bold)

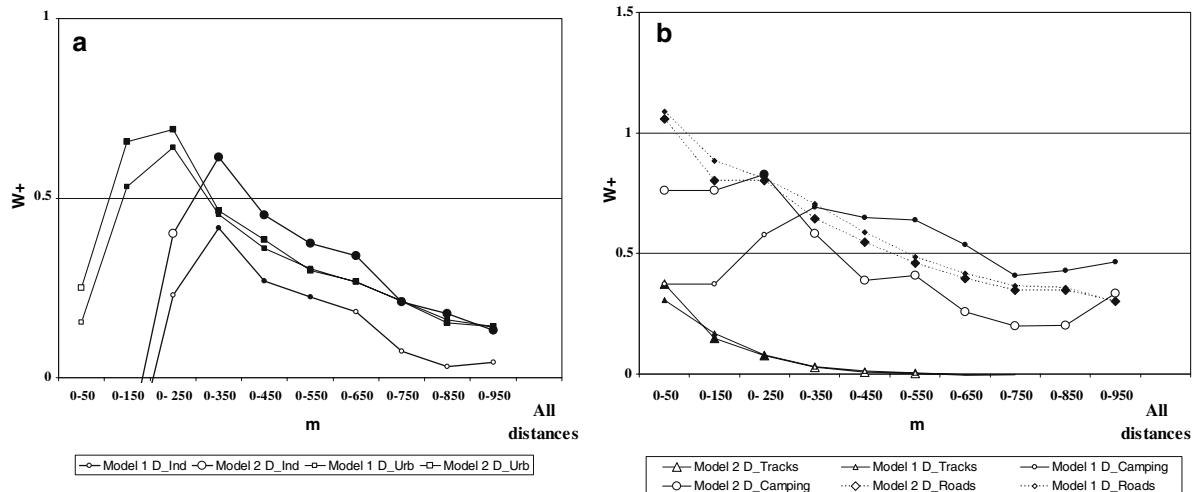


Fig. 2 W^+ (a) Distance to Urban Areas and industrial areas (m) and (b) Distance to roads, tracks and camping areas (m) for 2000–2003. C_5 serves as a guide to the significance of the

with a certain distance (<550 m) from the industrial areas in both models.

With regards to Distance to Roads (Fig. 2b), the results show significant values for all classes. The high W^+ indicates a strongly spatial association and statistically significant ($\alpha = 0.05$). The W^+ values for variable in the cumulative class 0–450 m are >0.5, indicating moderately predictive values. In relation to Distance from Tracks, the result shows that ignition risk is less in areas distant from tracks (>250 m). A distance <50 m has the most influence on ignition. The Distance to Camping Areas shows moderate predictive power in the cumulative class 0–250 (Model 1) and the cumulative classes from up to 250 to 650 m (Model 2).

The comparisons of result between the Model 1 and Model 2 show distinct similarities. The importance of Distance to Roads and Density of Sheep does not differ between the Models and do not show differences greater than $\pm 15\%$ with respect to Model 1. Densities of Population, Density of Secondary Housing, Distance of Urban Areas, Density of Cattle and Distance to Tracks show minor differences. Only one of the classes, normally the first class in the distance layers, show differences greater than $\pm 15\%$ with respect Model 1. Four of the 11 variables show differences in the majority of the classes between the Models. The Density of goats, Distance to Camping areas,

spatial association. C_5 values >1.96 indicate that the hypothesis that $C = 0$ can be rejected at $\alpha = 0.05$ (these significant values are shown by filled symbols)

Distance to Recreational Areas, Distance to Industrial Areas show different behavior between the two and four month models. Especially interesting are the Distance to Recreational Areas and Distance to Camping. In that layer, the predictive classes are different between the models and only some of them are statistically significant ($\alpha = 0.05$). For example, for Distance to Camping Areas all the high values of W^+ are from up to 250 to 950 m. in Model 1. The most predictive range is 0–350 m. However, the results of Model 2 show significant differences to Model 1. The results from Model 2 show that it is only statistically significant ($\alpha = 0.05$) at range 0–250 m. In this range, the values are $W^+ = -0.8$, $W^- = -0.02$, $C_5 = -2.1$. The value of W^+ denotes a higher predictability than Model 1 in this range. The Distance-buffer from Recreational Areas analysis shows that wildfire ignition occurs in areas closer to recreational areas. Only the range 0–4500 m of Model 1 shows statistical significance ($\alpha = 0.05$), but the value of W^+ suggest low predictability ($W^+ = 0.15$). Bonham-Carter et al. (1989) indicated that if C_5 is >1.96 the value of C is statistically significant ($\alpha = 0.05$), and indicates high confidence. Methodologically it is important to note how the WofE method allows the rejection of these layers from the modelling process at an early stage. However, a good understanding of the variables and its reason for rejection from the model is needed.

Generalizing the evidence maps and testing the conditional independence assumption

The results of the generalization of the evidence maps (Table 4) are useful to analyze and understand the weight of each layer in relation to the ignition. All layers considered have an influence on wildfire ignition and are statistically significant ($\alpha = 0.05$). However, not all have the same influence. Only five and six variables in Model 1 and Model 2, respectively, have a Contrast >0.7. Four of the variables are the same in both models, but Distance to Industrial Areas, Distance to Tracks and Distance to Camping Areas have different contrasts. Distance to Industrial Areas and Distance to Camping Areas increase in influence when the number of months considered is reduced to two. This highlights the importance of indicating the different behavior of the evidence maps between the models. Six of the 11 maps show differences greater than $\pm 15\%$ with respect Model 1.

The statistical validity of the resulting predictive maps is examined by considering a contingency table based on all pairs of maps, using a chi-squared test. For Model 1, five layers were selected; for Model 2 four layers were selected (Table 4). The results of the Overall Test of Conditional Independence were 0.96 for Model 1 and 1.00 for Model 2. Thus, the Pairwise Conditional Independence Test and Overall Test of Conditional Independence suggest that our predictive maps are statistically valid.

Creating a predictive map of human-based ignition risk

We show two of the possible predictive maps of Human-Based Ignition Risk (Fig. 3) as the ratio of the posterior probability to the prior probability. They have similar areas of each risk class for both periods (Table 5). The ‘high’ ignition risk class occupies 9.0% and 9.9% of the study area for maps produced by Models 1 and 2 respectively. The ‘low’ ignition risk class occupied 56% in Model 1 and 58% in Model 2. The values of the AAF show that in the case of ‘low’ ignition risk the number of wildfire ignitions observed is less than half of that expected. In the case of ‘high’ ignition risk, twice the number of wildfire ignitions occurred than would be expected by chance.

The spatial patterns of ignition risk vary between predictive maps. In the map from Model 1, the spatial pattern is defined mainly by distance to roads. In the Model 2 map, the distance to an urban area has most of the influence in defining the spatial pattern.

Model validation

The predictive model was compared with ignition points observed in 2004, 2005 and 2004–2005 (Table 5) showing excellent results. The accuracy of the Model 1 is similar between known occurrences and testing data in term of proportion of ignition point and AAF. The AAF is quite similar between the

Table 4 Contrast and confidence of evidence maps

Abbreviation	Model 1: 4 months		Model 2: 2 months		Base100 Model 1 (Model 2 Contrast) * 100/(Model 1 Contrast)
	Contrast	Confidence	Contrast	Confidence	
Dy_Pop	0.60	4.45	0.66	4.09	109.96
Dy_2Housing	0.51	3.52	0.48	2.61	95.65
Dy_Cattle	0.74	5.79	0.87	5.61	116.66
Dy_Sheep	0.96	7.26	0.92	5.80	96.21
Dy_Goats	0.45	3.47	0.58	3.72	128.70
D_Recreat	0.35	2.69			
D_Urb	1.08	4.79	1.24	4.51	115.48
D_Ind	0.54	1.99	1.96	2.69	364.70
D_Roads	0.85	5.84	0.94	5.29	111.31
D_Tracks	0.73	3.29	0.62	2.34	84.57
D_Camping	0.53	3.07	0.85	2.19	159.47

Evidence maps shown in bold are used to build the Predictive Map in each model. Those values in italic and bold indicate differences greater than $\pm 15\%$ between the models with respect to Model 1

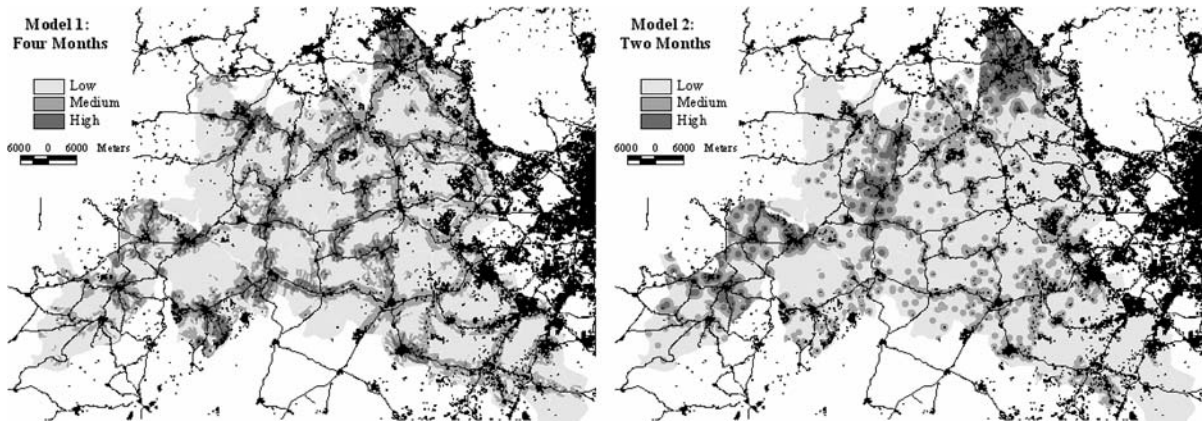


Fig. 3 Predictive maps of human-based ignition risk

Table 5 Results of models and validation from the human-based ignition risk

Model 1. Four months							Model 2. Two months						
	ha	%	Known occurrences					ha	%	Known occurrences			
			IP	%	P/ha	AAF				IP	%	P/ha	AAF
Low	100,920.9	55.8	74	29.8	4.5	0.53	Low	99,549.9	57.7	51	30.2	5.1	0.53
Medium	63,523.0	35.1	121	48.8	13.5	1.39	Medium	56,068.3	32.5	85	50.3	15.2	1.57
High	16,315.4	9.0	53	21.4	22.7	2.37	High	17,007.4	9.9	33	19.5	19.4	1.88
Validate 2004–2005							Validate 2004–2005						
	ha	%	Testing data					ha	%	Testing data			
			IP	%	P/ha	AAF				IP	%	P/ha	AAF
Low	100,920.9	55.8	45	26.8	4.5	0.48	Low	99,549.9	57.7	23	22.8	2.3	0.39
Medium	63,523.0	35.1	86	51.2	13.5	1.46	Medium	56,068.3	32.5	38	37.6	6.8	1.16
High	16,315.4	9.0	37	22.0	22.7	2.44	High	17,007.4	9.9	40	39.6	23.5	4.02
Validate 2004							Validate 2004						
	ha	%	IP	%	P/ha	AAF		ha	%	IP	%	P/ha	AAF
Low	100,920.9	55.8	24	25.8	2.4	0.46	Low	99,549.9	57.7	13	25.5	1.3	0.44
Medium	63,523.0	35.1	48	51.6	7.6	1.47	Medium	56,068.3	32.5	24	47.1	4.3	1.45
High	16,315.4	9.0	21	22.6	12.9	2.50	High	17,007.4	9.9	14	27.5	8.2	2.79
Validate 2005							Validate 2005						
	ha	%	IP	%	P/ha	AAF		ha	%	IP	%	P/ha	AAF
Low	100,920.9	55.8	21	28.0	2.1	0.50	Low	99,549.9	57.7	10	20.0	1.0	0.35
Medium	63,523.0	35.1	38	50.7	6.0	1.44	Medium	56,068.3	32.5	14	28.0	2.5	0.86
High	16,315.4	9.0	16	21.3	9.8	2.36	High	17,007.4	9.9	26	52.0	15.3	5.28

IP: Ignition point; P/ha: Density of ignition points * 10,000; AAF: Area-adjusted frequencies (Predicted/Expected ratio)

known occurrences and the result of testing data. However, we found that results using test data for Model 2 are better than the known occurrences. In

Model 2, the prediction map discriminates a higher number of ignition point in the high ignition risk class.

Discussion and conclusion

Socio-economic characteristics and dynamics

Because humans cause the majority of ignitions in the study area, the measures of accessibility (Roads, Path, Building, etc.) are important descriptors of the effects of the ubiquitous human population. The new lifestyles that are beginning to be led in the study area are characterized by increased recreation time and activities and the spread of urbanization into forest areas, in conjunction with increased human mobility to distant forest areas. These changes in behavior are the main factors driving the spatial distribution of people in forest areas and of the increase in ignition events (Badia-Perpinyà and Pallares-Barbera 2006; Venevsky et al. 2002). These changes indicate a potential shift in the nature of Spanish wildfire risk, as lifestyles and landscapes become similar to those found in other Mediterranean regions of the world with stronger economies, such as California. For example, Syphard et al. (2007a, b) also found that fire frequency in California, where mean distance to low-density housing has decreased recently, was well modelled by factors such as population density and distance to the wildland-urban interface (WUI).

The urbanization of rural areas has increased the WUI. Further, population expansion and recreational activities has increased ignition risk. Accordingly, Distance to Urban and Building Areas is the most predictive variable in our ignition risk model (Table 4). Prestemon et al. (2002) and Badia-Perpinyà and Pallares-Barbera (2006) found similar results, identifying the WUI as a statistically significant wildfire risk factor. As others have suggested (Prestemon et al. 2002; Butry et al. 2001), we believe that the risk of economic damage from wildfire in the WUI might be sufficient incentive to further refine our understanding of the relationship between wildfire and human factor. However, as Syphard et al. (2007a, b) highlight, understanding how the WUI itself influences wildfire is in itself an important question for ecologists. Integrated socio-ecological and ecological-economic models of succession-disturbance dynamics are now being developed and are likely to become an important area of research in the future (Perry and Millington 2007).

We have found, as have others (Badia-Perpinyà and Pallares-Barbera 2006; de Vasconcelos et al.

2001), that the spatial patterns of ignition are strongly associated with landscape accessibility. The weights of the classes (Table 3, Fig. 2) and the contrast of evidence maps (Table 4) confirm this. As ours, Cardille et al. (2001), Chou et al. (1993), de Vasconcelos et al. (2001), Badia-Perpinyà and Pallares-Barbera (2006) and Dickson et al. (2006) have found that the proximity to roads and tracks is positively correlated with ignition. As with Yang et al. (2007), we have found that <50 m is a key distance and <450 m is the threshold and upper occurrence limit. Chuvieco and Salas (1996) also concluded that distance to Roads, Recreational areas and Trails was an important predictor of an ignition risk in Sierra de Gredos (Central Spain).

In our study area, the presence of goats and sheep has a direct relationship with absence of wildfire ignition (Table 3). It is likely that the abandonment of traditional livestock farming and the change to new uses (e.g. recreational) is involved in ignition increases. A medium density of cattle is a good predictor of ignition. However, the discontinuity of the cattle density classes could be a consequence of effect of level of data aggregation.

We identified significant interactions between the same variables within the two-month and four-month wildfire ignition risk models. However, our results show that the ignition regime differs depending on the four-month (Model 1) and two-month (Model 2) season. Distances from industrial areas have an influence on ignition risk, and is more important in Model 2 than Model 1 (Fig. 2). Results for Distance to Camping Areas indicates that wildfire ignition occurs closer to these areas and that there are significant spatial and temporal patterns. In Model 2, the proximity to these areas has more influence than Model 1 on the ignition risk. However, the Distance to Recreation Areas is not found to have a significant influence on wildfire ignition location. This may be because the mobility of people is not adequately represented by the defined buffer classes and because that information is based on a point layer. This aspect of the modeling requires further investigation.

As Cardille et al. (2001) found, all the socio-economic and spatial variables (except Distance to Recreational Area in Model 2) used in our analysis are correlated with ignition risk. Our results (Table 4) emphasize distinct patterns regarding regional human-based ignition fires in the study area and

reflect the most important socio-economic variables in different season periods (Model 1 and Model 2).

The predictive maps (Fig. 3) predict rural municipalities are at greater risk than metropolitan municipalities within the study area. In Model 2, the higher risk areas are situated on the SW-NE axis, where there is low density urbanization near areas with aesthetic value and where recreational resources are greater. These characteristics are likely to attract more people for their summer holidays in July and August. Our predictive maps included the majority of evidence maps with high contrast values. Whilst the accuracy of the model is excellent, our results are influenced by the variables included in each predictive map. Using other evidence maps may result in other spatial patterns of predicted risk. Because of this, we think that it is necessary to explore the other possible combinations to ensure our findings are robust. This aspect requires further research.

We have found that our analysis is useful for examining different aspects of fire risk and for assessing the usefulness of variables included in the model. WofE provides a quantitative tool to relate socio-economic processes with spatial patterns of wildfire ignitions. WofE has proven to be a useful approach here because it explicitly considers the spatial association between ignition occurrence and evidence map data, and is relatively straight-forward to implement and interpret the results (e.g. Dickson et al. (2006) and Romero-Calcerrada and Luque (2006)). Our results, together with Dickson et al. (2006), suggests WofE analysis could be used to spatially predict and analyze wildfire ignition. The weights estimated from WofE analysis are useful for evaluating impacts of different variables on ignition risk. The prediction map from Model 1 is more robust statistically than the prediction map from Model 2, because the AAF show similar validation results (Table 5). However, these results also suggest that the predictive capacity of the WofE approach is robust to small samples sizes at large/medium scales (Model 2 has a smaller sample size).

Fire management planning implications

Fire management planning, particularly for the long term, requires an understanding of the relationships between spatial patterns and causes of wildfire

human-caused ignition risk. These relationships can be investigated from many perspectives. We suggest that models such as that presented here (i.e. WofE) will help to identify in space and time the significant socio-economic factors causing ignition and in turn will be useful to optimize time, human resources and fire management funds. The wildfire ignition risk maps produced here will be useful in the spatially explicit assessment of fire risk, for combination with models such as FARSITE, the planning and coordination of regional efforts to identify areas at greatest risk, and for designing large-scale wildfire management strategies.

The present study advances the understanding of the spatial dynamics of human-caused wildfire occurrence. One overall implication of our research is that wildfire models estimated for rural regions could be improved by including socio-economic factors in addition to biophysical variables (e.g. fuel type, fuel moisture content, temperature, etc.). Including the socio-economic factors could clarify this cartography and define different class of risk. In this way, Integrated Risk Models can help in the design of more effective and efficient wildland fire management and public policy (Mercer and Prestemon 2005). Few studies have sought to assess socio-economic factors that could be affecting wildfire ignition risk. Consistent with Prestemon et al. (2002) we believe that the understanding of fire risk must include socio-economic variables and patterns of human activity.

Acknowledgements We thank the two anonymous reviewers for the extremely valuable comments provided as well as the editorial comments. We would like to express our gratitude to Servicio de Cartografía Regional and Sección de Defensa Contra Incendios Forestales of the Regional Government of Madrid for the Digital cartography and Ignition Point database. We also wish to thank at the Ministry of Education and Sciences for the finance received (project CGL2004-06049-C04-02/CLI, FIREMAP).

References

- Agterberg FP, Bonham-Carter GF, Wright DF (1990) Statistical pattern integration for mineral exploration. In: Gaal G, Merriam DF (eds) Computer applications in resource estimation. Pergamon Press, Oxford, pp 1–21
- Aspinall R (1992) An inductive modelling procedure based on Bayes' theorem for analysis of pattern in spatial data. *Int J GIS* 6:105–121

- Aspinall R, Veitch N (1993) Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modeling procedure in a GIS. *Photogramm Eng Remote Sens* 59:537–543
- Badia-Perpinyà A, Pallares-Barbera M (2006) Spatial distribution of ignitions in Mediterranean periurban and rural areas: the case of Catalonia. *Int J Wildland Fire* 15: 187–196
- Bonham-Carter GF (1994) *Geographic information systems for geoscientists, modelling with GIS*. Pergamon, Tarrytown, New York
- Bonham-Carter GF, Agterberg FP, Wright DF (1989) Weights of evidence modelling: a new approach to mapping mineral potential. In: Agterberg FP, Bonham-Carter GF (eds) *Statistical applications in the earth science*. Geological Survey of Canada, pp 171–183
- Bonham-Carter GF, Agterberg FP (1999) Arc-WofE: a GIS tool for statistical integration of mineral exploration datasets. The 52 Session of the International Statistical Institute. *Bulletin of the International Statistical Institute*, Helsinki, Finland, p 4
- Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FKA (2002) Evaluating resource selection functions. *Ecol Modell* 157:281–300
- Butry DT, Mercer DE, Prestemon JR, Pye JM, Holmes TP (2001) What is the price of catastrophic wildfire?. *J Forestry* 99:9–17
- Cardille JA, Ventura SJ, Turner MG (2001) Environmental and social factors influencing wildfires in the upper Midwest, United States. *Ecol Appl* 11:111–127
- Carranza EJM, Hale M (2000) Geologically Constrained probabilistic mapping of gold potential, Baguio District, Philippines. *Nat Resour Res* 9:237–253
- Carranza EJM, Hale M (2002) Where are porphyry copper deposits spatially localized? A case study in Benguet Province, Philippines. *Nat Resour Res* 11:45–59
- Chou YH, Minnich RA, Salazar LA, Power JD, Dezzani RJ (1990) Spatial autocorrelation of wildfire distribution in the Idyllwild Quadrangle, San Jacinto Mountain, California. *Photogramm Eng Remote Sens* 56:1507–1513
- Chou YH, Minnich RA, Chase RA (1993) Mapping Probability of fire occurrence in San Jacinto Mountains, California, USA. *Environ Manage* 17:129–140
- Chuvieco E, Congalton RG (1989) Application of remote-sensing and GIS to forest fire hazard mapping. *Remote Sens Environ* 29:147–159
- Chuvieco E, Salas J (1996) Mapping the spatial distribution of forest fire danger using GIS. *Int J Geogr Inf Syst* 10: 333–345
- D. G. del Medio Natural (1997) *Mapa de Vegetación y Ocupación del Suelo*. Consejería de Medio Ambiente y Ordenación del Territorio. Comunidad de Madrid, Madrid
- D. G. del Medio Natural (2000) *Áreas Recreativas (Cartografía)*. Consejería de Medio Ambiente y Ordenación del Territorio. Comunidad de Madrid, Madrid
- de Vasconcelos MJP, Silva S, Tome M, Alvim M, Pereira JMC (2001) Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. *Photogramm Eng Remote Sens* 67:73–81
- Dickson BG, Prather JW, Xu YG, Hampton HM, Aumack EN, Sisk TD (2006) Mapping the probability of large fire occurrence in northern Arizona, USA. *Landsc Ecol* 21:747–761
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Modell* 135:147–186
- Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecol Modell* 199:142–152
- Hooten MB, Larsen DR, Wikle CK (2003) Predicting the spatial distribution of ground flora on large domains using a hierarchical Bayesian model. *Landsc Ecol* 18:487–502
- INE (1999) *Censo Agrario*. Instituto Nacional de Estadística, Madrid
- INE (2001) *Censo de Población y Vivienda*. Instituto Nacional de Estadística, Madrid
- Kemp LD, Bonham-Carter GF, Raines GL (1999) WofE: Arcview extension for weights of evidence mapping. <http://www.ntsर्व.gis.nrcan.gc.ca/wofe>
- Kemp LD, Bonham-Carter GF, Raines GL, Looney CG (2001) Arc-SDM: arcview extension for spatial data modelling using weights of evidence, logistic regression, fuzzy logic and neuronal networks analysis. <http://www.ntsर्व.gis.nrcan.gc.ca/sdm/>.
- Mercer DE, Prestemon JP (2005) Comparing production function models for wildfire risk analysis in the wildland-urban interface. *Forest Policy Econ* 7:782–795
- Milne BT, Johnston KM, Forman RTT (1989) Scale-dependent proximity of wildlife habitat in a spatially-neutral Bayesian model. *Landsc Ecol* 2:101–110
- Millington JDA (2005) Wildfire risk mapping: considering environmental change in space and time. *J Mediterr Ecol* 6:33–42
- MMA (2007) *Los incendios forestales en España. Decenio 1996–2005. Área de Defensa Contra Incendios Forestales*. Ministerio de Medio Ambiente, Madrid, p 106
- Moreno JM, Vázquez A, Vélez R (1998) Recent history of forest fires in Spain. In: Moreno JM (ed) *Large forest fires*. Backhuys, Leiden, pp 159–185
- Pearce JL, Boyce MS (2006) Modelling distribution and abundance with presence-only data. *J Appl Ecol* 43: 405–412
- Perry GLW, Millington JDA (2007) Spatial modelling of succession-disturbance dynamics in terrestrial ecological systems. *Perspect Plant Ecol Evol Systematics* doi:10.1016/j.ppees.2007.07.001
- Prestemon JP, Pye JM, Butry DT, Holmes TP, Mercer DE (2002) Understanding broadscale wildfire risks in a human-dominated landscape. *For Sci* 48:685–693
- Romero-Calcerrada R, Perry GLW (2002) Landscape change pattern (1984–1999) and implications for fire incidence in

- the SPA Encinares del rio Alberche y Cofio (Central Spain). In: Viegas DX (ed) Forest fire research & wildland fire safety. Millpress, Rotterdam
- Romero-Calcerrada R, Perry GLW (2004) The role of land abandonment in landscape dynamics in the SPA 'Encinares del rio Alberche y Cofio, Central Spain, 1984–1999. *Landsc Urban Plan* 66:217–232
- Romero-Calcerrada R, Luque S (2006) Habitat quality assessment using weights-of-evidence based GIS modeling: the case of *Picoides tridactylus* as species indicator of the biodiversity value of the Finnish forest. *Ecol Modell* 196:62–76
- SCR (2000) Mapa Topográfico. Consejería de Obras Públicas, Urbanismo y Transportes. D. G. de Planificación Territorial. Comunidad de Madrid, Madrid
- Skidmore AK (1989) An expert system classifies eucalypt forest types using Thematic Mapper data and a digital terrain model. *Photogramm Eng Remote Sens* 55: 1449–1464
- Syphard AD, Clarke KC, Franklin J (2007a) Simulating fire frequency and urban growth in southern California coastal shrublands, USA. *Landsc Ecol* 22:431–445
- Syphard AD, Radeloff VC, Keely JE, Hawbaker RJ, Clayton MK, Stewart SI, Hammer RB (2007b) Human influence on California Fire Regimes. *Ecol Appl* 17:1388–1402
- Tucker K, Rushton SP, Sanderson RA, Martin EB, Blaiklock J (1997) Modelling bird distributions—a combined GIS and Bayesian rule-based approach. *Landsc Ecol* 12:77–93
- Vázquez A, Moreno JA (1995) Patterns of fire occurrence across a climate gradient and its relationship to meteorological variables in Spain. In: Moreno JA, Oechel WC (eds) *Global change and Mediterranean-type ecosystems*. Springer Verlag, New York, pp 408–434
- Venevsky S, Thonicke K, Sitch S, Cramer W (2002) Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study. *Glob Chang Biol* 8:984–998
- WWF/Adena (2004) Incendios Forestales. Causas, situación actual y propuestas. WWF/Adena, Madrid, p 25
- WWF/Adena (2005) Incendios Forestales. ¿Porqué se queman los bosques españoles? WWF/Adena, Madrid, p 49
- Yang J, Healy HS, Shifley SR, Gustafson EJ (2007) Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. *For Sci* 53:1–15