Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/forestsystems

Forest Systems 2011 20(1), 95-107
ISSN: 1131-7965
eISSN: 2171-9845

Modelling the probability of lightning-induced forest fire occurrence in the province of León (NW Spain)

F. Castedo-Dorado[1]*, J. R. Rodriguez-Perez[2], J. L. Marcos-Menendez[3]
and M. F. Alvarez-Taboada

[1] Departamento de Ingeniería y Ciencias Agrarias. Escuela Superior y Técnica de Ingeniería Agraria. Campus de Ponferrada. Universidad de León. Avda. de Astorga, s/n. 24400 Ponferrada (León). Spain
[2] Departamento de Tecnología Minera, Topográfica y de Estructuras. Campus de Ponferrada. Universidad de León. Avda. de Astorga, s/n. 24400 Ponferrada (León). Spain
[3] Departamento de Química y Física Aplicadas. Universidad de León. Escuela de Ingenierías Industrial e Informática. Campus Vegazana, s/n. 24071 León. Spain

Abstract

Spatial relationships between lightning-induced forest fires and topography, vegetation, climate and lightning characteristics were analyzed in the province of León (NW Spain). The study was based on reported lightning-induced forest fires in the period 2002-2007. A statistical model based on logistic regression was developed to estimate the probability of occurrence of a lightning-induced fire in a 3 × 3 km grid. The importance of accurate location of the ignition point was also investigated in order to evaluate the sensitivity of the model developed to uncertainty of the location. The model developed with accurate ignition point data showed a better predictive ability than the model constructed with all the ignition points available. The former model was therefore selected for long-term prediction of the occurrence of lightning-induced fires in the province. According to this model, the probability of a forest stand being affected by lightning-induced fire increased with decreasing altitude, and when there was a high proportion of coniferous species in the stand, a high percentage of lightning strikes in forest areas and a high number of dry storm days in the area.

Although the model has not been validated, the results can be considered spatially robust because it shows good classification ability and the predicted spatial probability distribution is consistent with the observed historical fire records. The model will be useful in the spatially explicit assessment of fire risk, the planning and coordination of regional efforts to identify areas at greatest risk, and in designing long-term wildfire management strategies.

Key words: lightning; forest fires; logistic regression; ignition probability; geographic information systems.

Resumen

Modelización de la probabilidad de ocurrencia de incendios forestales por rayo en la provincia de León (NO de España)

En este estudio se analizaron las relaciones espaciales entre incendios forestales originados por rayo y variables topográficas, de vegetación, de clima y de las características de las descargas de rayos en la provincia de León (NO de España). Se utilizaron datos de incendios forestales originados por rayo en la provincia en el período 2002-2007. La probabilidad de ocurrencia de incendio se estimó para una malla de 3 km × 3 km de lado mediante un modelo logístico. Se analizó también la importancia que posee una localización exacta del punto de inicio del incendio en los resultados del modelo desarrollado y la sensibilidad del mismo a la incertidumbre de la localización. El modelo desarrollado a partir de todos los puntos de inicio de incendio disponibles en la base de datos mostró una peor capacidad predictiva que el desarrollado a partir de datos exactos del punto de inicio. Este último fue, por tanto, el seleccionado para realizar la predicción a largo plazo de los incendios provocados por rayo en la provincia. De acuerdo con este modelo, la probabilidad de ocurrencia de un incendio causado por rayo aumenta según disminuye la altitud, cuando existe una elevada proporción de coníferas, un elevado porcentaje de descargas de rayos en terreno forestal y un elevado número de días de tormentas secas.

Aunque el modelo no ha sido validado, los resultados del mismo pueden considerarse espacialmente robustos ya que el modelo tiene una buena capacidad de clasificación y la distribución espacial predicha es consistente con los

---

* Corresponding author: fcasd@unileon.es
Received: 12-07-10; Accepted: 28-01-11.

registros históricos de incendios. El modelo obtenido es útil en la valoración explícita del riesgo de incendios, en la identificación de zonas de riesgo elevado, así como en el diseño de estrategias para la gestión de incendios forestales a largo plazo.

Palabras clave: rayos; incendios forestales; regresión logística; probabilidad de ignición; sistemas de información geográfica.

Introduction

Fire is the most serious threat to Spanish forests. Lightning is the major natural cause of forest fire ignition in Spain and throughout the world (Pyne *et al.,* 1996), and is the main overall cause of ignition in some areas, such as boreal forest. Human-induced fires are predominant in countries around the Mediterranean Basin and in other similar areas around the world (Vázquez and Moreno, 1998), and because of the importance of such fires in the current forest fire regimes of these countries, little attention has been given to lightning-induced fires. However, lightning-ignited fires may burn larger areas of forest than human-induced fires because of their remoteness and aggregation in time and space (Podur *et al.,* 2003).

Several authors have reported that lightning-induced fires do not occur at random, but rather tend to start in specific places (Vankat, 1985). The efficiency of individual lightning strikes in igniting a forest fire is affected by variation in lightning properties such as quantity, polarity and intensity; moisture properties of forest fuel resulting from recent weather conditions including precipitation, temperature and humidity; topographic variables that may affect the above-mentioned variables (*e.g.* Díaz-Ávalos *et al.,* 2001), and rates of combustion, which vary with type of fuel. The relative importance of all these variables varies with the scale considered, and it is unrealistic to present a general model for large scales, thus making it advisable to develop models at local or regional scales (Pacheco *et al.,* 2009).

In the province of León (NW Spain), lightning-induced fires accounted for 5% of the total area burned during the period 1983-2006 (and constituted approximately 5% of the total wildfires during this period). However, in some years (*e.g.* 1996 and 2005) natural wildfires led to more than 20% of the area in the province being burned. In addition, the statistical records show an increase in lightning-induced fires (especially in number of fires and fire days) during the last ten years of the study period (Fig. 1). Similar increases have also been reported in other studies in Spain (*e.g.* Vázquez and Moreno, 1998) and are difficult to explain, since changes in methods of fire detection and record keeping cannot be ruled out (Vázquez and Moreno, 1998).

The aim of the present study was to identify the significant factors related to lightning-induced forest fire occurrence (risk) and to model their spatial pattern the province of León. The final objective is to include the relevant factors in a GIS in order to provide forest managers with a long-term probability map for lightning-induced forest fires.

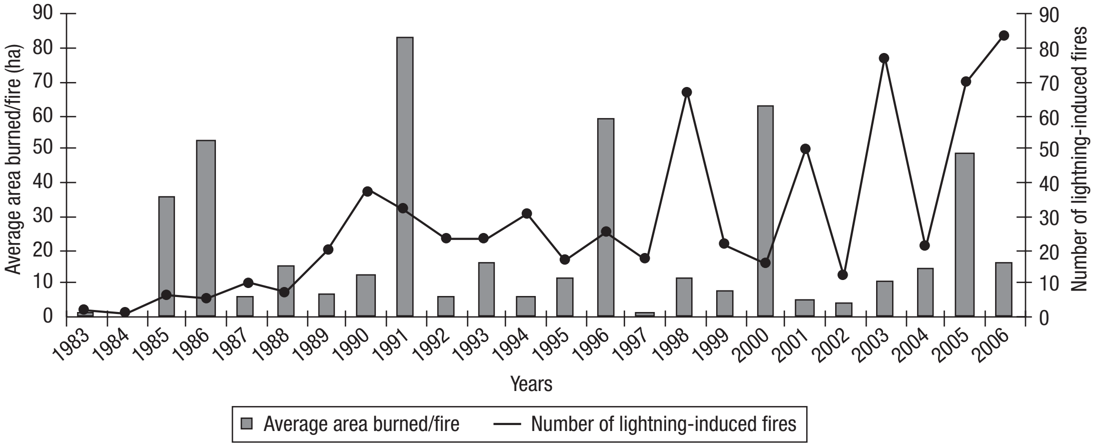

Figure 1. Changes in the average area burned per lightning-induced fire and of the total number of lightning-induced fires in the province of León for the period 1983-2006.

Material and methods

Data

The data used to develop the lightning fire occurrence model were obtained from five different sources. The study comprises the months of May to September (inclusive) in a six year period (2002-2007). The five month interval was considered because both lightning-caused fires and lightning strokes occurred almost exclusively during these months (Fig. 2). The particular period (2002-2007) was selected because: (i) 2002 was the first year that some wildfire ignition points were recorded in the province (*i.e.,* using X,Y coordinates); (ii) it enables the assumption that land cover does not differ from that in 2003, the year for which land cover data was available for the present study; (iii) flash detection efficiency has improved greatly in Spain in recent years since the detection network began to operate in 1992. Since 2001 the detection efficiency has exceeded 90%, and average location accuracy is within 0.5 km (Pérez-Puebla, 2004).

The study area was partitioned in pixels of 3 × 3 km, resulting in a total of 1,882 pixels for the province of León. In addition, all digital information available for the analysis was converted to this 3 × 3 km spatial resolution by use of ArcGis 9.2 software by ESRI. This particular grid size was selected because it represents an adequate compromise between resolution requirements, interpolation accuracy and computation cost. A larger pixel size would lead to loss of information about multiple ignition occurrences, whereas a smaller pixel size would make it difficult to detect any spatial pattern in the data. Other studies on this subject have used the same or very similar pixel sizes (*e.g.* Díaz-Ávalos *et al.,* 2001; Nieto *et al.,* 2006).

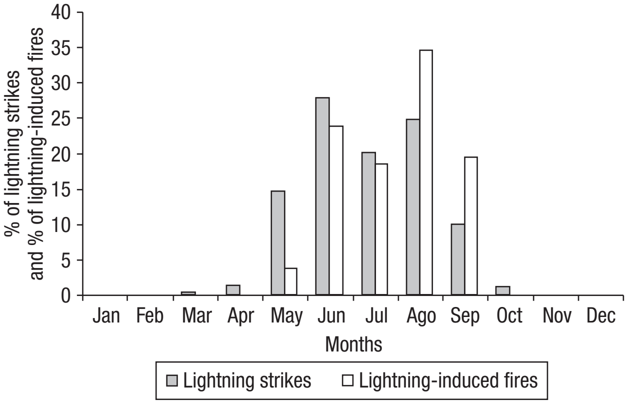

Figure 2. Monthly distribution of the percentage of lightning strikes and lightning-induced fires for the period 2002-2007 in the province of León.

Lightning location data was provided by the Spanish State Meteorological Agency (Agencia Estatal de Meteorología: AEMET). The database includes information such as lightning intensity, polarity, date (to nearest second), estimated coordinates of lightning strikes and quality of location estimation. According to Martín-León (1999) and Nieto *et al.* (2006), only accurately located flashes (with χ^2 equal or smaller than 2 and a long semi-axis radius of the error ellipse equal or smaller than 1.5 km) were selected. The database finally used included 78,256 flashes, with less than 5% of negative polarity. A similar percentage was reported by Rivas-Soriano *et al.* (2005) for the province of León, in a study on the lightning activity in the Iberian Peninsula in the period 1992-2005.

The ignition locations of lightning-induced forest fires were derived from a fire database generated by the Spanish Ministry of the Environment and Rural and Marine Affairs. This database provides, among other things, information about detection time, ignition location and estimated cause of ignition. Some fire records do not include the UTM coordinates (X,Y) of the ignition points, but only the name of the local entity or the municipality and the 10 × 10 km UTM grid where the wildfire started. According to the information available, three data sets were distinguished: (1) lightning-induced fires for which the UTM coordinates were known, (2) lightning-induced fires for which the UTM coordinates or the name of the local entity were known, and (3) all lightning-induced fires reported, *i.e.,* fires in data sets (1) and (2) and those for which only the name of the municipality and the 10 × 10 km UTM grid were known. If the exact point where the fire started was not recorded, the criteria for assignation to a 3 × 3 km pixel were based on spatial intersections of local entity or Municipality and the 10 × 10 km UTM grid, and on the daily observations of the spatial distribution of lightning strokes.

In order to obtain the topographic variables, a digital terrain model (DTM) with a pixel resolution of 50 m provided by the National Cartographic Database was used. From this DTM, the average slope, altitude and aspect were calculated and assigned to each pixel of 3 × 3 km, using the Spatial Analysis module of ArcGIS 9.2 software by ESRI.

Information about the composition and structure of land cover (scale 1:50,000) was obtained from the digital Spanish Forest Map of the province of León

(Spanish Ministry of the Environment, 2003). In order to deal with more homogeneous types of land cover, the information provided by the Spanish Forest Map was categorized in the following classes: coniferous woodland, broadleaf woodland, mixed woodland, non combustible areas, gallery woodland, open woodland, mosaic composed of woodland and others, shrubland, grassland, recently deforested areas, and recent forest plantations and reforestations.

Daily meteorological data for a 3 × 3 km grid was derived from raw data recorded at 344 weather stations located in León and nearby provinces. This data was provided by AEMET, and consists of hourly or daily observations of temperature, relative humidity and rainfall. Geostatistical methods were used to represent these data in continuous surfaces. Universal kriging was used as an interpolation method for rainfall and relative humidity data. This method models the spatial autocorrelation, assuming that the trend varies depending on the data position (on the contrary, ordinary kriging assumes that the trend is constant). The co-kriging interpolation method was used to model temperature as the main variable of interest and elevation as a secondary variable. This method is recommended when there is correlation between a secondary variable (in this study, elevation) and the variable of interest (Hudson and Wackernagel, 1994). The use of elevation as a secondary variable has been found to improve the spatial variation in daily climatological variables even when a slight correlation with elevation was observed (*e.g.* Jarvis and Stuart, 2001; Carrera-Hernández and Gaskin, 2007).

Methodology

The spatial patterns of occurrence were analyzed by searching for relationships between the presence/absence of lightning-caused forest fires in a 3 × 3 km grid and the potential explanatory variables converted to this scale by use of a geographic information system.

In order to transform the meteorological variables (relative humidity and temperature) into structural variables, the number of days throughout 2002-2007 when the observed values were above a defined threshold, were used. The thresholds were computed as the mean values recorded during thunderstorms that caused forest fires.

In addition, the number of dry storms in the period was counted as the number of thunderstorms with an accumulated precipitation in 24 hours equal to or less than 2.5 mm. This threshold is based on previous studies on lightning-induced fires (Rorig and Ferguson, 1999; Álvarez-Lamata, 2001; Rorig and Ferguson, 2002).

The 1882 pixels in which the province of León was divided were coded with 0 (absence) or 1 (occurrence of one or more lightning-induced ignitions during the period 2002-2007). Three alternative codifications were computed, according to the three different data sets considered (see Data Section).

Among the many distribution functions that have been proposed for use in the analysis of this type of presence-absence binary dependent variable, logistic distribution is recommended because of its flexibility and widespread use. Logistic regression analysis was therefore used to model the probability of occurrence of lightning-induced forest fires in any 3 × 3 km prediction unit (pixel). Logistic regression analysis has been used successfully for predicting fire occurrence and for examining the most critical factors involved in fire incidence (Vega-García *et al.,* 1995; Vasconcelos *et al.,* 2001; Martínez *et al.,* 2009).

Variable selection

One of the most important stages in developing this kind of model is to select the best set of independent variables to explain the probability of lightning-induced ignitions. The most relevant factors reported in the relevant literature were initially included as independent variables in the general logistic model.

In order to find which variables were the most significant, a standard stepwise logistic regression analysis was carried out using the SAS/STAT® LOGISTIC procedure (SAS Institute Inc., 2004). The significance level for entering variables into the model and retaining them was set at 0.05. This criterion for variable selection was complemented by attempting to include non related variables to avoid multicollinearity, and searching for a logical interpretation of the resulting parameter estimates.

Model evaluation and validation

Measurements of goodness-of-fit for dichotomous variables take into account that it does not matter how close a prediction for a pixel of 3 × 3 km is to 0 or to 1, as long as it classifies the observation correctly. The goodness-of-fit of the compared models was measured using the value of the generalization of the coefficient of determination (R^2) proposed by Cox and Snell (1989) and modified by Nagelkerke (1991), and the Hosmer-Lemeshow goodness-of-fit test (χ^2_{HL}) (Hosmer and Lemeshow, 2000). Large values of the latter statistics (and small p-values) indicate a lack of a good model fit (SAS Institute Inc., 2004).

The area under the receiver operating characteristic (ROC) curve was also calculated using the fitting dataset. This curve relies on false/true-positive/negative tests, where sensitivity is the proportion of event responses that were predicted to be events and specificity is the proportion of non-event responses that were predicted to be non-events (SAS Institute Inc., 2004). The plot of sensitivity (*i.e.,* hit rate) *versus* 1-specificity (*i.e.,* false alarm rate) is the ROC curve; the area under this curve measures the accuracy of the detection system and does not require any assumptions concerning the shape or form or the underlying signal and noise distributions (Saveland and Neuenschwander, 1990). This statistic is a threshold-independent measure of model discrimination, where 0.5 suggests no discrimination, 0.7-0.8 suggests acceptable discrimination, and 0.8-0.9 suggests excellent discrimination (Hosmer and Lemeshow, 2000).

A cut-off approach can be used to convert the continuous probabilities produced by the lightning-induced fire occurrence model to dichotomous results (*i.e.,* presence or absence). This method specifies a fixed cut-off (c) or threshold between 0 and 1. If the estimated ignition probability was less than c, the outcome value was «absence»; otherwise it was «presence». In this study, we compared two different methods of establishing this threshold. In the first method, the threshold was considered as the average observed proportion of positive ignition pixels *versus* total pixels in the province. In the second method, the cut-off point was established by computing and representing the sensitivity and specificity for different cut-off values. The optimal cut-off point corresponds to the value where the sensitivity curve and specificity curve cross (Hosmer and Lemeshow, 2000).

The correct classification rate (CCR) was calculated to assess the performance of both methods. This was based on a classification table of correct/incorrect responses for prediction of lightning fire occurrence in each of the 3 × 3 km pixels into which the province was divided. In addition, the accuracy of the classification was measured by its sensitivity and specificity.

A portion of the data was not reserved for model validation. The only method that can be regarded as «true» validation involves the use of a new independent data set (Vanclay and Skovsgaard, 1997; Yang *et al.,* 2004), although the scarcity of such data obliges the use of alternative approaches. The common method of splitting the data set in two portions does not provide additional information (*e.g.,* Huang *et al.,* 2003) and is not recommended from the point of view of parameter estimation (Myers, 1990; Hirsch, 1991). Moreover, other techniques such as double cross-validation or statistical tests provide very limited information about the predictive ability of the models (Kozak and Kozak, 2003; Yang *et al.,* 2004).

Results

The explanatory variables selected for estimating the probability of occurrence of lightning-induced fires are shown in Table 1. Depending on the data set analyzed (*i.e.,* the three cases considered according to the uncertainty of ignition location), the variables included varied slightly, although altitude, percentage of strokes in forest area and number of dry thunderstorm days were significant for all three data sets, indicating that these are the most relevant factors in the occurrence of lightning-induced fires in the province of León.

The likelihood ratio test revealed that all parameters were significant indicators of lightning-induced fire (Table 1). The values of the generalized coefficient of determination were 0.159, 0.175 and 0.138, for data sets 1, 2 and 3, respectively. Although these values do not appear to be very adequate, it must be taken into account that the R^2 for logistic models is not analogous to the R^2 in OLS regression (*i.e.,* it cannot be interpreted in terms of explained variation). Similar values were obtained, for example, by González *et al.* (2006) and Chuvieco et al. (2010) in developing fire risk models for Catalonia (NE Spain) and other regions in Spain, respectively.

The chi-square values and associated probability for the Hosmer and Lemeshow (2000) goodness-of-fit test (χ^2_{HL}) indicates that there was no evidence of any significant difference between predicted and observed survival probability at $\alpha = 0.05$ for data sets 1 and 2. For data set 3, this test indicates that there is a lack of fit for the model selected.

The area under the ROC curve (Fig. 3) indicated excellent discrimination for data sets 1 and 2 (area = 0.797

Table 1. Parameter estimates and their standard errors, and goodness-of-fit statistics of the equations for predicting lightning-induced fires in the province of León.

Variable	Parameter estimate	Standard error	P > Chi-Sq.	% concord. pairs	R^2	χ^2_{HL}	P > Chi-Sq.
Data set 1 (86 fires; 75 pixels)							
Intercept	–5.13	0.715	< 0.0001	80.2	0.159	12.52	0.1295
Altitude	–0.00211	0.000540	< 0.0001				
% coniferous woodland area	0.0225	0.00563	< 0.0001				
% strokes in forest area	0.0396	0.00756	< 0.0001				
Number of dry thunderstorm days	0.137	0.0371	0.0002				
Data set 2 (111 fires; 98 pixels)							
Intercept	–4.01	0.556	< 0.0001	79.3	0.175	9.966	0.2674
Altitude	–0.00257	0.000484	< 0.0001				
% strokes in forest area	0.0352	0.00575	< 0.0001				
% strokes in coniferous woodland area	0.0166	0.00520	0.0014				
Number of dry thunderstorm days	0.155	0.0334	< 0.0001				
Data set 3 (298 fires; 179 pixels)							
Intercept	–3.49	0.414	< 0.0001	73.8	0.138	15.80	0.0453
Altitude	–0.00218	0.000418	< 0.0001				
% woodland area	0.0118	0.00417	0.0047				
% strokes in forest area	0.0153	0.00419	0.0003				
Number of dry thunderstorm days	0.143	0.0261	< 0.0001				
Number of less-than-average-moisture days	0.00214	0.000709	0.0026				

and 0.792, respectively). For data set 3, the discrimination can be considered acceptable (area = 0.741). At this point, it must be recognized that use of the same data for model fitting and for calculating the area under the ROC curve, may lead to an over-optimistic estimate of the performance, which the score would provide if it were to be validated on a sample of future cases (Copas and Corbett, 2002).

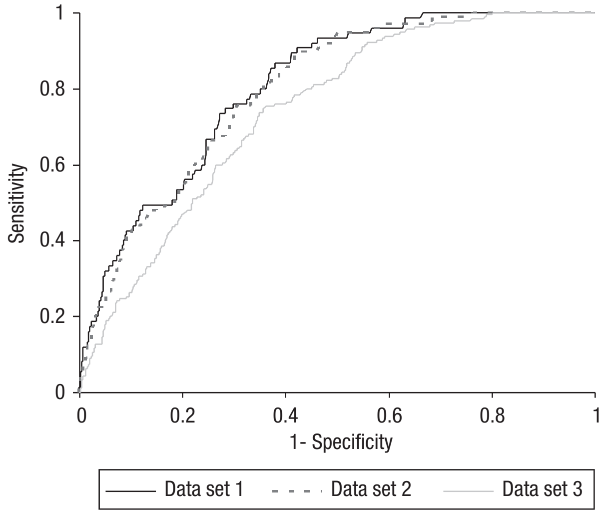

Figure 3. ROC curves for the probability of occurrence of lightning-induced fires in the province of León for the three data set used.

Parameter estimates (Table 1) were used to calculate the probability of lightning-induced fire for each 3 × 3 km pixel, and for the three data sets analyzed. Therefore, the two different methods of cut-off determination produced six different thresholds for comparing the estimated and the observed status (presence or absence of a lightning-induced fire) of each pixel (Table 2).

According to the first method of cut-off determination, the thresholds were established at 3.98%, 5.20% and 9.51% for data sets 1, 2 and 3, respectively, since these were the mean observed proportions of positive ignition pixels in relation to total pixels, for the province. Alternatively, with the method in which the sensitivity and the specificity curves are considered (see Fig. 4), the threshold values were 4.4%, 5.5% and 10.7% for data sets 1, 2 and 3, respectively. It can therefore be concluded that both methods provided similar cut-off points for the three data sets analyzed.

The cut-off values obtained using both alternative methods provided a similar percentage of pixels with presence of lightning-induced fires. However, these values were very different from the observed values (33.4%, 34.6% and 41.6% for the first method of cut-

Table 2. Accuracy of classification of positive/negative ignition, considering both the observed percentage of positive ignition pixels and the point at which the sensitivity and specificity curves crossed as cut-off values

Data set	Cuf-off value (*c*)	Sensitivity (%)	Specificity (%)	Correct classification rate (CCR, %)	Predicted positive ignition pixels (%)
1	0.0398	76.0	68.3	68.6	33.4
	0.044	74.7	71.2	71.4	30.5
2	0.0052	74.5	67.5	67.9	34.6
	0.055	73.5	69.2	69.4	33.0
3	0.0951	75.9	61.9	63.2	41.6
	0.107	67.6	67.2	67.2	36.1

off determination, and 30.5%, 33% and 36.1% for the second method).

The CCR was in the interval 63-71% for all data sets, and was slightly higher when the cut-off value was calculated using the sensitivity and specificity curves. The percentage of correctly classified pixels was larger when the ignition point of the lightning-induced fires was collected accurately (*i.e.,* data set 1). On the other hand, the sensitivity values were always higher than specificity values and, as expected, the specificity increased and the sensitivity decreased as the cut-off value increased.

Probability estimates from data set 1 (*i.e.,* those related with ignition locations most accurate collected) and a cut-off value of 0.044 provided the highest CCR, and were therefore selected in order to determine the presence/absence of lightning-induced fires in a 3 × 3 km grid.

Plots of the observed altitude, percentage of coniferous woodland area, percentage of lightning strokes in forest area and number of dry thunderstorm days against observed and predicted probability of the occurrence of lightning-induced fires show that the models provide good predictions of the occurrence of these fires (Fig. 5). Probability estimates from data set 1 was used in this figure.

The location of the observed and predicted lightning-induced fires in the province of León for the period 2002-2007 is shown in Figure 6. Only the 75 pixels corresponding to data set 1 were represented for the observed fires. For representation of the predicted lightning-caused fires, probability estimates from data set 1 and a cut-off value of 0.044 were used. The model estimated a significant percentage of false positives, *i.e.,* prediction of the occurrence of lightning-caused fires, which in fact, did not occur.

Discussion

According to the goodness-of-fit statistic used for evaluation, the logistic models developed in this study were found to be adequate for estimating lightning-induced fire occurrence in the province of León. The percentage of pixels for occurrence of a correctly classified lightning fire ranged between 68% and 76%, and the percentage of pixels for non occurrence of a lightning ranged between 62% and 71%, depending on

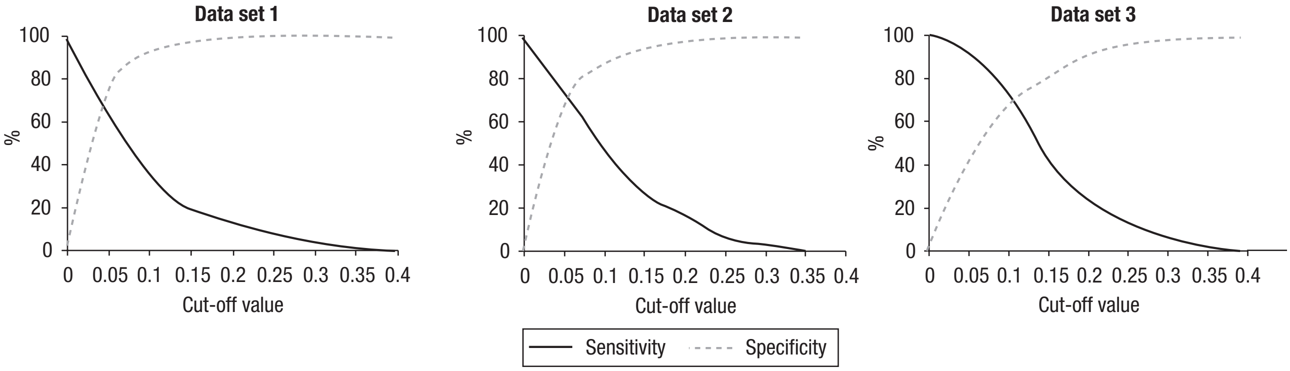

Figure 4. Sensitivity and specificity of the models developed for the three data sets concerning lightning-induced fires analyzed.

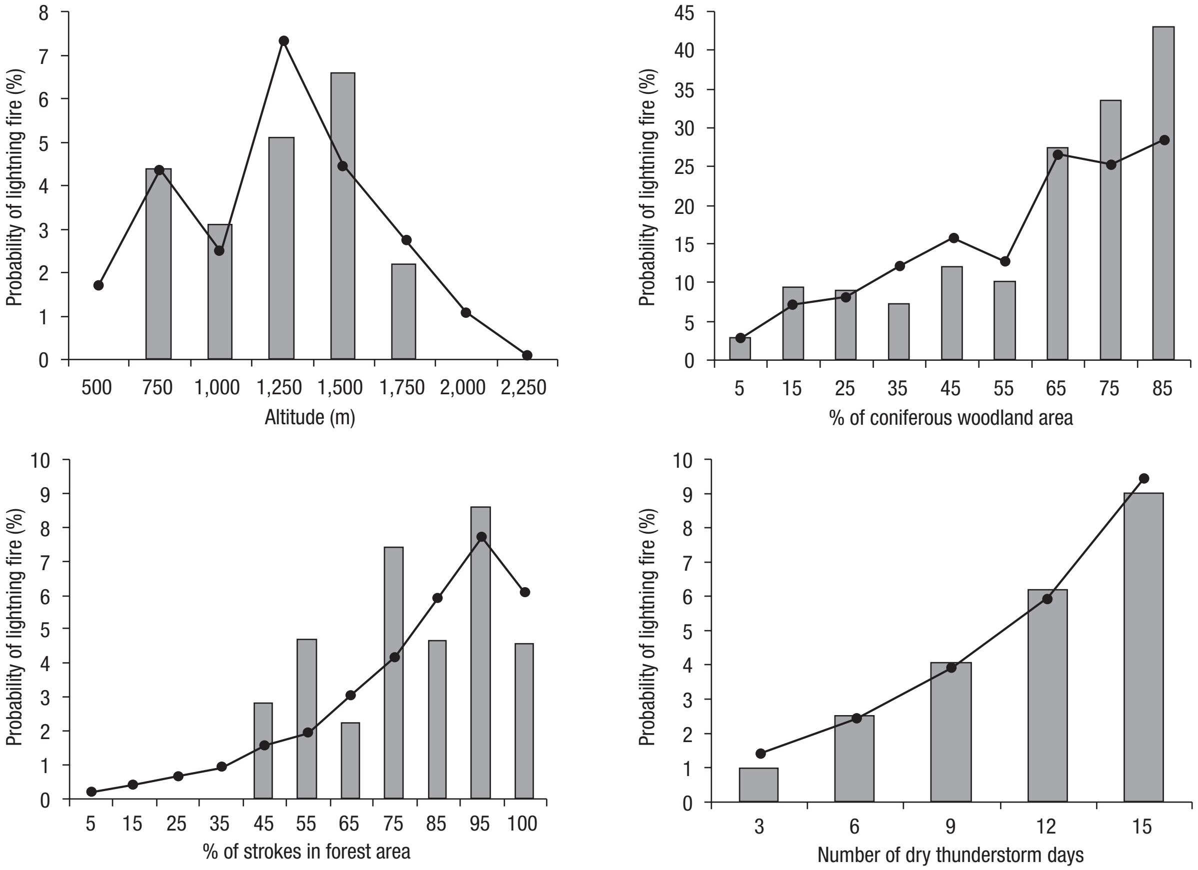

Figure 5. Predicted (line) and observed (bar) occurrences of lightning-induced fires for the explained variables obtained in the regression analysis from data set 1.

the data set analyzed and the cut-off used. The values of the CCR obtained are somewhat larger than those obtained in some studies on the occurrence lightning-induced fires in other regions of Spain (*e.g.,* Nieto *et al.,* 2006; Pacheco *et al.,* 2009; Chuvieco *et al.,* 2010).

Although the results reveal that the model displayed a good capacity for classification ability, the predictive

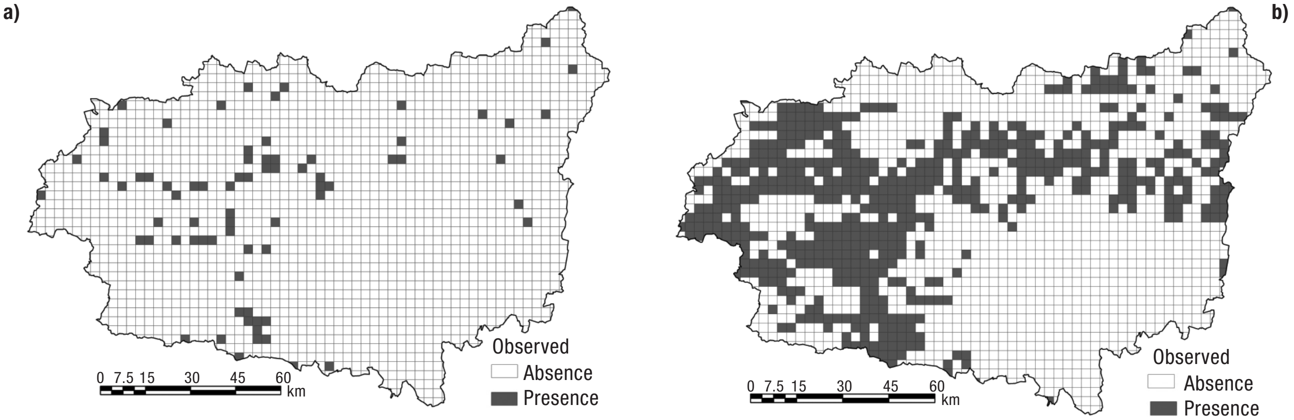

Figure 6. Location of the observed (a) and predicted (b) lightning-induced fires in the province of León for the period 2002-2007. Pixels corresponding to data set 1 were represented for the observed fires. Probability estimates from a logistic model (fitted using data set 1) and a cut-off value of 0.044 were used to represent the pixels in terms of presence/absence of lightning-induced fires.

capacity was not so good. Using as cut-off values both the observed percentage of positive ignition pixels and the point at which the sensitivity and specificity curves crossed, the percentage of pixels where a lightning-caused fire will occur is over-predicted. However, it must be taken into account that the number of pixels where lightning fires do not occur is much greater than the percentage of pixels where lighting fires do occur, so that errors resulting in overestimation of the number of pixels in relation to fire occurrence are of much less importance.

As expected, the classification and predictive ability of the models developed from data sets in which location of ignition points is affected by a high degree of uncertainty was lower than those in which this is negligible. This highlights the importance of the exact location of the point of ignition for developing feasible models of wildfire occurrence (Amatulli *et al.,* 2007).

The results shown in Table 1 demonstrate that altitude had a negative effect on the probability of occurrence of lightning-induced fires, *i.e.,* lower and intermediate elevations were found to be the most prone to fire. Although several studies have shown a positive relation between lightning occurrence and altitude (*e.g.* Dissing and Verbyla, 2003), even in the study region (Rivas Soriano *et al.,* 2001*,* 2005), higher rainfall and lower temperatures at higher altitudes may cause a negative relationship (Martín, 1982; Díaz-Ávalos *et al.,* 2001). Moreover, the altitudinal ecological limit of woodlands may also be important (Dissing and Verbyla, 2003). These outcomes suggest that lightning-induced fires occur at altitudes where fuel continuity and moisture are not limiting factors (Martin, 1982). Some other authors have reported similar results for lightning-induced fires in other regions of Spain (*e.g.,* Nieto *et al.,* 2006).

No relationship was found between the other topographical variables (slope and aspect) and the occurrence of lightning-induced fires. Similar results were found in other studies (*e.g.,* by McRae, 1992, in Australia). Other authors have reported that fires occurring on terrain facing suntraps are more likely to spread due to higher solar incidence and comparatively drier fuel (Vasconcelos *et al.,* 2001). Conedera *et al.* (2006), Wierzchowski *et al.* (2002) and Díaz-Ávalos *et al.* (2001) found that lightning-induced fires mainly occur on steeper slopes.

The presence of forest areas and woodlands is positively associated with lightning-caused fire ignitions. These results are consistent with those of Vázquez and Moreno (1998), who found that lightning-induced fires in Spain affected a greater proportion of woodlands than human-induced fires. This may be the result of the canopy sheltering the forest floor from rainfall associated with lightning (Kourtz and Todd, 1992).

More surprising, *a priori*, is the significance of variables related to the existence of coniferous woodlands (percentage of coniferous woodland area in data set 1 and percentage of strokes in coniferous woodland area in data set 2), taking into account that this type of vegetation represents less than 25% of the woodland area in the province of León (Junta de Castilla y León, 2005). This confirms that some types of vegetation cover are more prone to lightning-induced fire than others (Manry and Knight, 1986; Granstrom, 1993; Dissing and Verbyla, 2003; Krawchuk *et al.,* 2006; Evett *et al.,* 2008).

Lightning ignition may occur when the electrical current ignites the fine fuels on the forest floor (usually the duff) at the base of a tree (Latham and Williams, 2001) or when living trees act as lightning conductors (Ogilcie, 1989). Differences in duff layer, produced by differences in vegetation type, would also result in different rates of heating and therefore differences in flammability (Latham and Williams, 2001). The duff layer of needles sheltered under conifers makes a more suitable location for a fire to start than the duff layer in a hardwood stand (Flannigan and Wotton, 1991). Deciduous trees species decrease the duff depth and consequently, the probability of ignition (Latham and Schlieter, 1989).

In addition, the high flammability of coniferous species (due to their high content of resin and essential oils) and the abundance of highly flammable species in the understory of the *Pinus* stands in the province (*Erica* sp., *Genistella tridentata*, *Calluna vulgaris*, etc.) may help fire propagation after ignition (*e.g.* Vélez, 1990; Bond and Van Wilgen, 1996).

Although some authors (*e.g.,* Flannigan and Wotton, 1991) have found that dead fuel moisture content (DFMC) is one of the best variables to explain the incidence of lightning-induced fire, the preference for a more operational model and the uncertainty associated with the estimation of fuel moisture have led to this variable being omitted, and therefore it was not included in the analysis in this study.

Meteorological danger indices (which rely on current and past weather conditions) have commonly been used to estimate DFMC. According to Aguado *et al.* (2007), there are two main difficulties associated with meteorological danger indices in DFMC estimation: spatial significance and calibration. The former is caused by the location of weather stations, which may not be the most appropriate for fire danger estimation, since they are commonly associated with monitoring agricultural or urban parameters. Consequently, the required spatial interpolation techniques always introduce a certain estimation error, which is added to the actual estimation of DFMC. Another operational difficulty in using meteorological danger indices to estimate DFMC conditions is the lack of calibration of moisture indices to ecosystem or climate characteristics other than where they were developed. Furthermore, the moisture content in the forest floor can vary considerably over very short distances. Even when a lightning strike occurs quite close to a weather station (where estimation of the local value of fuel moisture would be most precise), the actual moisture content at the exact location of the discharge through the forest floor may be considerably different (Wotton and Martell, 2005).

Occurrence of dry thunderstorm is also one of the most important variables in the probability of lightning-induced ignition for the three data sets analyzed, and is highly significant in all of them. Some authors have noted that dry thunderstorms (defined as thunderstorms without significant concurrent rainfall) are common and an important contributor to fire ignition. A lightning fire is more likely to be observed when precipitation on the day of a lightning strike is null or negligible, possibly because greater precipitation extinguished fires prior to discovery (Hall, 2007; Rorig *et al.,* 2007). In addition, it must be taken into account that rainfall affects fuel moisture content.

Logistic regression analysis has not shown any relation between lightning density and fuel ignition. This may be because higher densities of flashes are followed by larger amounts of precipitation (Álvarez-Lamata, 2005; Rorig and Ferguson, 1999; Rorig and Ferguson, 2002). There is no obvious relationship between lightning strike density and lightning-induced fire density (Podur *et al.,* 2003; Larjavaara *et al.,* 2005; Hall, 2007, and also confirmed for Spain by Vázquez and Moreno, 1993). This suggests that, for the province of León, there does need to be a very large number of lightning strikes for several ignitions to take place.

There was also no evident relationship between the occurrence of lightning-induced fires and the main characteristics of lightning discharges: charge (positive or negative) and intensity. Positive lightning occurrence was not related to the temporal occurrence of fires. This contradicts the findings of several studies in which it was concluded that positive lightning is more likely to cause ignition because there is a greater likelihood of a long continuum current (Latham and Williams, 2001; Anderson, 2002; Wotton and Martell, 2005), and confirms the findings of Flannigan and Wotton (1991) and Larjavaara *et al.* (2005). This observed lack of correlations suggest that ignition coincides more closely with available fuels than with the number, charge or some other lightning strike characteristics (Rorig and Ferguson, 1999).

The other meteorological variables considered (temperature and relative humidity) were not significant for the two data sets in which the most accurate data was obtained (data sets 1 and 2). The arbitrary definition of the thresholds for the indices that allow conversion of these meteorological variables in structural variables, and their correlation with altitude may explain the lack of significance.

The resulting map of the probability of occurrence of lightning-induced fire in the period 2002-2007 is shown in Figure 7. Probability estimates were obtained from the model developed with data set 1, since it provides the highest CCR. As reported in Results, absence of lightning-induced fires is expected for pixels in which the probability of fire occurrence is less than 0.044. As can be observed, the highest probability of fire is concentrated in areas of intermediate altitude in the province, where woodlands are the dominant type

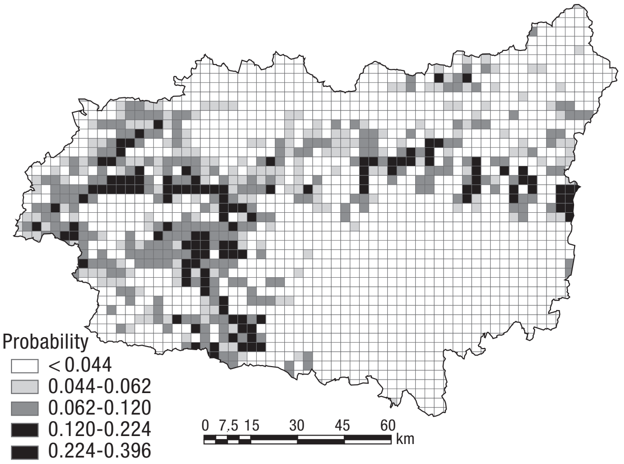

Figure 7. Spatial distribution of the probability of occurrence of lightning-induced fires in the province of León for the 2002-2007 period. Probabilities were based on the logistic model and parameter estimates obtained from data set 1. Absence of lightning-induced fires is expected for pixels in which the probability of fire occurrence is less than 0.044.

of land cover. This result suggests, as Amatulli et al. (2007) pointed out, that lightning-induced fires are concentrated in particular areas. This clustered pattern may be explained by local vegetation and weather conditions, which directly influence the fuel moisture content of the different types of vegetation.

Conclusions

A model for estimating the probability of occurrence of lightning-induced fires in the province of León is presented. Logistic regression techniques were used to generate predictive models at 3 km × 3 km spatial resolution.

The variables selected are not the same in all three data sets considered, although the variables typically related to fire ignition and propagation (topography, vegetation and weather conditions) are important in all three. Altitude, percentage of strikes in forest area and occurrence of dry thunderstorms were found to be common variables for the three data sets analyzed. The importance of all these variables has also been recognized by the fire-research community.

Potential sources of variability in these models are the interpolation of weather-based observations, errors in estimating lightning position and, above all, the errors in estimating lightning-induced fire ignitions. The model developed from accurately identified ignition points was found to have a better classification and predictive ability than the model that included all the ignition points available. The former model was therefore selected for long-term prediction of the occurrence of lightning-induced fire in the province. This result also highlights the importance of the exact location of the point of ignition for developing feasible models of the occurrence of lightning-induced fires.

Outputs from this model can be considered robust and are therefore good indicators for assessing the risk of naturally occurring forest fires and a useful tool for forest fire prevention planning in the province of León. Longer observation periods will enable future improvements in the model.

Risk maps for lightning-induced ignition will be useful in the spatially explicit assessment of fire risk, for combination with models such as *FARSITE* (Finney, 2004), the planning and coordination of regional efforts to identify areas at greatest risk, and for designing long-term wildfire management strategies.

In addition, the model can easily be integrated with other sources of risk (*e.g.* socio-economic causes), within the structure of geographic information systems. The results can also provide a framework for understanding how future changes in landscape features (*e.g.,* percentage of coniferous species) may influence the occurrence of lightning-induced fires.

Acknowledgements

Funding for this research was provided by the Diputación of León within the Project «Modelización de la probabilidad espacial y temporal de ocurrencia de incendios forestales por rayo en la provincia de León». The AEMET provided the data on lightning strikes and the meteorological data used.

References

AGUADO I., CHUVIECO E., BORÉN R., NIETO H., 2007. Estimation of dead fuel moisture content from meteorological data in Mediterranean areas. Applications in fire danger assessment. Int J Wild Fire 16, 390-397.

ÁLVAREZ-LAMATA E., 2001. Factor de tormentas. Proc V Simposio Nacional de Predicción. Sección C: Técnicas y herramientas de análisis, diagnosis y predicción. Madrid, Spain. Nov. 20-23. 6 pp. [In Spanish].

ÁLVAREZ-LAMATA E., 2005. Los incendios forestales y las condiciones meteorológicas en Aragón. Proc 4º Congreso Forestal Español, Zaragoza, Spain. Sept. pp. 26-30. 7 pp. [In Spanish].

AMATULLI G., PÉREZ-CABELLO F., DE LA RIVA FERNÁNDEZ J., 2007. Mapping lightning/human-caused fires occurrence under ignition point location uncertainty. Ecol Model 200, 321-333.

ANDERSON K., 2002. A model to predict lightning-caused fire occurrences. Int J Wild Fire 11, 163-172.

BOND W.J., VAN WILGEN B.W., 1996. Why and how do ecosystems burn? In: Fire and plants. Ed Chapman & Hall, London. pp. 17-33.

CARRERA-HERNÁNDEZ J.J., GASKIN S.J. 2007. Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. J. Hydrol 336, 231-249.

CHUVIECO E., AGUADO I., YEBRA M., NIETO H., SALAS J., MARTÍN M.P., VILAR L., MARTÍNEZ J., MARTÍN S., IBARRA P., DE LA RIVA J., BAEZA M.J., RODRÍGUEZ F., MOLINA J.R., HERRERA M.A., ZAMORA R., 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. Ecol Model 221, 46-58.

COPAS J.B., CORBETT P., 2002. Overestimation of the receiver operating characteristic curve for logistic regression. Biometrika 89, 315-331.

CONEDERA M., CESTI G., PEZZATTI G.B., ZUMBRUNNEN T., SPINEDI F., 2006. Lightning-induced fires in the Alpine region: an increasing problem. Proc 5th International conference on forest fire research. Coimbra, Portugal. Nov. 27-30.

COX D.R., SNELL E.J., 1989. Analysis of binary data, 2nd ed. Chapman and Hall, London.

DÍAZ-ÁVALOS C., PETERSON D.L., ALVARADO E., FERGUSON S.A., BESAG J.E., 2001. Space-time modelling of lightning-caused ignitions in the Blue Mountains, Oregon. Can J For Res 31, 1579-1593.

DISSING D., VERBYLA D.L., 2003. Spatial patterns of lightning strikes in interior Alaska and their relations to elevation and vegetation. Can J For Res 33, 770-782.

EVETT R.R., MOHRLE C.R., HALL B.L., BROWN T.J., STEPHENS S.L., 2008. The effect of monsoonal atmospheric moisture on lightning fire ignitions in southwestern North America. Agr For Meteo 148, 1478-1487.

FINNEY M.A., 2004. *FARSITE*: Fire Area Simulator-model development and evaluation. USDA Forest Service Res. Pap. RMRS-RP-4, Ogden, UT. 47 pp.

FLANNIGAN M.D., WOTTON B.M., 1991. Lightning-ignited forest fires in Northwestern Ontario. Can J For Res 21, 277-287.

GONZÁLEZ J.R., PALAHÍ M., TRASOBARES A., PUKKALA T., 2006. A fire probability model for forest stands in Catalonia (north-east Spain). Ann For Sci 63, 169-176.

GRANSTROM A., 1993. Spatial and temporal variation in lightning ignitions in Sweden. J Veg Sci 4, 737-744.

HALL B.L., 2007. Precipitation associated with lightning-ignited wildfires in Arizona and New Mexico. Int J Wild Fire 16, 242-254.

HIRSCH R.P., 1991. Validation samples. Biometrics 47, 1193 1194.

HOSMER D., LEMESHOW S., 2000. Applied logistic regression. Wiley-Interscience, New York. 392 pp.

HUANG S., YANG Y., WANG Y., 2003. A critical look at procedures for validating growth and yield models. In: Modelling forest systems (Amaro A., Reed D., Soares P., eds). Ed CAB International, Wallingford, Oxfordshire, UK. pp. 271-293.

HUDSON G., WACKERNAGEL H., 1994. Mapping temperature using Kriging with external drift: theory and example from Scotland. Int J Clim 14, 77-91.

JARVIS C.H., STUART N., 2001. A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: the interaction between number of guiding variables and the type of interpolation method. J Appl Meteorol 40, 1075-1084.

JUNTA DE CASTILLA Y LEÓN, 2005. Castilla y León crece con el bosque. Consejería de Medio Ambiente. Serie Divulgativa. 48 pp. [In Spanish].

KOURTZ P.H., TODD B., 1992. Predicting the daily occurrence of lightning-caused forest Fires. Inf Rep PI-X-112. Canadian Forest Service, Petawawa.

KOZAK A., KOZAK R., 2003. Does cross validation provide additional information in the evaluation of regression models? Can J For Res 33, 976-987.

KRAWCHUK M.A., CUMMING S.G., FLANNIGAN M.D., WEIN R.W., 2006. Biotic and abiotic regulation of lightning fire initiation in the mixedwood boreal forest. Ecology 87, 458-468.

LARJAVAARA M., PENNANEN J., TUOMI T.J., 2005. Lightning that ignites forest fires in Finland. Agr For Meteor 132, 171-180.

LATHAM D., SCHLIETER J.A., 1989. Ignition probabilities of wildland fuels based on simulated lightning discharges, USDA Forest Service, Research Paper INT-411. Ogden, UT. 16 pp.

LATHAM D., WILLIAMS E., 2001. Lightning and forest fires. In: Forest fires, behavior and ecological effects (Johnson E.A., Miyanishi K. eds). Academic Press. pp. 375-418.

MANRY D.E., KNIGHT R.S., 1986. Lightning density and burning frequency in South African vegetation. Vegetatio 66, 67-76.

MARTIN R.E., 1982. Fire history and its role in succession. In: Forest succession and stand development research in the Northwest (Means J.E., ed). USDA Forest Service Forest Research Laboratory,, Oregon State University, Corvallis. 92-98.

MARTÍNEZ J., VEGA-GARCÍA C., CHUVIECO E., 2009. Human-caused wildfire risk rating for prevention planning in Spain. J. Environ Manage 90, 1241-1252.

MARTÍN-LEÓN F., 1999. Caracterización de la actividad tormentosa peninsular y áreas limítrofes durante el periodo estival de 1994. Proc IV Simposio Nacional de Predicción del I.N.M, Madrid, Spain, April 15-19 1996. [In Spanish].

MCRAE R., 1992. Prediction of areas prone to lightning ignition. Int J Wild Fire 2, 123-130.

MYERS RH., 1990. Classical and modern regression with applications. Boston, USA, PWS-KENT Publishing Company. 488 pp.

NAGELKERKE N.J.D., 1991. A note on a general definition of the coefficient of determination. Biometrika 78, 691-692.

NIETO H., AGUADO I., CHUVIECO E., 2006. Estimation of lightning-caused fires occurrence probability in central Spain. Proc 5th International conference on forest fire research. Coimbra, Portugal, Nov 27-30. 15 pp.

OGILCIE C.J., 1989. Lightning fires in Saskatchewan Forests. Fire Manage Notes 50, 31-32.

PACHECO C.E., AGUADO I., NIETO H., 2009. Análisis de ocurrencia de incendios forestales causados por rayo en la España peninsular. Geofocus 9, 232-249. [In Spanish].

PÉREZ-PUEBLA F., 2004. Cooperación entre las redes de rayos de España y Portugal. Proc Jornadas Científicas de la Asociación Meteorológica Española, Badajoz, Spain. Feb 11-13. 10 pp. [In Spanish].

PODUR J., MARTELL D.L., CSILLAG F., 2003. Spatial patterns of lightning-caused forest fires in Ontario, 1976-1998. Ecol Model 164, 1-20.

PYNE S.J., ANDREWS P.J., LAVEN R.D., 1996. Introduction to wildland fire. Second edition, John Wiley & Sons, New York-Chichester. 769 pp.

RIVAS SORIANO L., DE PABLO F., GARCÍA E., 2001. Cloud-to-ground lightning activity in the Iberian Peninsula: 1992-94. J Geophys Res 106, 11891-11901.

RIVAS SORIANO L., DE PABLO F., TOMÁS C., 2005. Ten-year study of cloud-to-ground lightning activity in the Iberian Peninsula. J Atmos Sol Ter Phys 67, 1632-1639.

RORIG M.L., FERGUSON S.A., 1999. Characteristics of lightning and wildland fire ignition in the Pacific Northwest. J Appl Meteor 38, 1565-1575.

RORIG M.L., FERGUSON S.A., 2002. The 2000 fire season: lightning-caused fires. J Atmo Sci 41, 786-791.

RORIG M.L., MCKAY S.J., FERGUSON S.A., WERTH P., 2007. Model-generated predictions of dry thunderstorm potential. J Appl Meteorol Clim 46, 605-614.

SAS INSTITUTE INC., 2004. SAS/STAT® 9.1User's Guide. SAS Institute Inc, Cary, NC.

SAVELAND J.M., NEUENSCHWANDER L.F., 1990. A signal detection framework to evaluate models of tree mortality following fire damage.