



# Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression

J. Martínez-Fernández<sup>1</sup>, E. Chuvieco<sup>2</sup>, and N. Koutsias<sup>3</sup>

<sup>1</sup>INIA-CIFOR, Forest Research Centre, Department of Forest Ecology and Genetics, Carretera de La Coruña, Km. 7,5 28040, Madrid, Spain

<sup>2</sup>Department of Geography and Geology, University of Alcalá, Calle Colegios 2, 28801 Alcalá de Henares, Spain

<sup>3</sup>Department of Environmental and Natural Resources Management, University of Western Greece, G. Seferi 2, 30100 Agrinio, Greece

Correspondence to: J. Martínez-Fernández (martinez.jesus@inia.es)

Received: 16 May 2012 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: –

Revised: 6 November 2012 – Accepted: 10 November 2012 – Published: 11 February 2013

**Abstract.** Humans are responsible for most forest fires in Europe, but anthropogenic factors behind these events are still poorly understood. We tried to identify the driving factors of human-caused fire occurrence in Spain by applying two different statistical approaches. Firstly, assuming stationary processes for the whole country, we created models based on multiple linear regression and binary logistic regression to find factors associated with fire density and fire presence, respectively. Secondly, we used geographically weighted regression (GWR) to better understand and explore the local and regional variations of those factors behind human-caused fire occurrence.

The number of human-caused fires occurring within a 25-yr period (1983–2007) was computed for each of the 7638 Spanish mainland municipalities, creating a binary variable (fire/no fire) to develop logistic models, and a continuous variable (fire density) to build standard linear regression models. A total of 383 657 fires were registered in the study dataset. The binary logistic model, which estimates the probability of having/not having a fire, successfully classified 76.4% of the total observations, while the ordinary least squares (OLS) regression model explained 53% of the variation of the fire density patterns (adjusted  $R^2 = 0.53$ ). Both approaches confirmed, in addition to forest and climatic variables, the importance of variables related with agrarian activities, land abandonment, rural population exodus and developmental processes as underlying factors of fire occurrence.

For the GWR approach, the explanatory power of the GW linear model for fire density using an adaptive bandwidth

increased from 53% to 67%, while for the GW logistic model the correctly classified observations improved only slightly, from 76.4% to 78.4%, but significantly according to the corrected Akaike Information Criterion ( $AIC_c$ ), from 3451.19 to 3321.19. The results from GWR indicated a significant spatial variation in the local parameter estimates for all the variables and an important reduction of the autocorrelation in the residuals of the GW linear model. Despite the fitting improvement of local models, GW regression, more than an alternative to “global” or traditional regression modelling, seems to be a valuable complement to explore the non-stationary relationships between the response variable and the explanatory variables. The synergy of global and local modelling provides insights into fire management and policy and helps further our understanding of the fire problem over large areas while at the same time recognizing its local character.

## 1 Introduction

Human factors are critical to explain fire occurrence worldwide, but they are particularly relevant in European Mediterranean countries with a long fire history related to traditional farming activities, as is the case in Spain. It is estimated that more than 90% of forest fires are caused by people in European Mediterranean countries (Leone et al., 2009; Vélez, 2009; FAO, 2007). Additionally, these areas have experienced important socioeconomic transformations over the last few decades, including land abandonment and/or higher

tourist and urban pressures on the forest areas, which could imply higher ignition risk. Given the importance of the human risk, any improvement in the modelling and assessment of factors that drive human-made ignitions is critical for fire prevention, planning and management. Also, a better knowledge of the spatial patterns of fire occurrence and their relationships with underlying factors of human risk becomes a necessity to locate and make prevention efforts more efficient.

It is also necessary to further improve the modelling techniques. In fire occurrence modelling, different statistical and regression modelling techniques have been applied at several temporal and spatial scales, in many cases assuming that the model parameters are valid and homogeneous for the entire study area from which the data were sampled or, alternatively, assuming that the model structure is spatially stationary, such as the examples in Syphard et al. (2008), Chuvieco et al. (2010), Vilar et al. (2010) and Kwak et al. (2012). However, when large geographical study areas are involved, it would be more reasonable to find varied rather than constant relationships. For instance, Koutsias et al. (2005, 2010), when modelling fire densities, observed that the explanatory power of classical linear regression increased considerably after assuming varying relationships instead of constant ones. Their analysis was developed at the provincial level (NUTS-3) across the European Mediterranean Basin countries (Portugal, Spain, southern France, Italy and Greece) using geographically seighted regression (GWR), so initiating the use of GWR in fire modelling studies.

Although interest in accounting for regional variations in wildfire occurrence factors has been shown recently in some studies (Moreira et al., 2009; Carmo et al., 2011; Gonzalez-Olabarria et al., 2011; Padilla and Vega-García, 2011; Nunes, 2012), except for Koutsias et al. (2005, 2010), this has only begun to be addressed very recently by using local geographically weighted regression (Tulbure et al. 2011; Poudyal et al., 2012; Avila-Flores et al., 2010; Sá et al., 2011; Rodrigues and De la Riva, 2012). In our study, similar to those of Koutsias et al. (2010) and Sá et al. (2011), GWR is considered as a complement to the “global” regression modelling approach, with which it is compared in order to better understand particular processes at the regional scale, but at the same time recognizing its own local characteristics and patterns (Fotheringham et al., 1996, 1997, 2002). The term “global” is used here to describe a model that refers to a homogeneous process in which the relationships being modelled are the same everywhere within the study area.

## 2 Objectives

The work presented here is an extension of previous research (Martínez et al., 2009) that showed how the rate of human-caused fires in Spain can be predicted and explained from socioeconomic and geographic variables, assuming spatially

stationary processes. The overall objective of this new study is to check, in a quantitative way, if these stationary models are adequate to properly explain and understand long-term fire occurrence patterns in a large study area such as the Spanish peninsular territory. In order to achieve this overall objective, three improvements have been implemented over the previous work.

The first concerns the predicted variable. Instead of modelling only the high versus low occurrence, in this paper we have addressed two aspects of fire occurrence: (i) fire presence/absence and (ii) fire density, using a longer historical time period (25 yr versus 13) for both. For these two aspects we built two predictive “global” models at the national scale using two “classical” regression approaches: OLS linear regression to explain long-term fire density patterns and, complementarily, a binary logistic model to define the existing underlying factors behind fire presence and to better understand why in some of the municipalities no fires have been observed during the studied period. The terms “ordinary” and “classical” are used here to represent the default regression model in many statistical software packages, in contrast to other specific models like GWR.

The second innovative aspect of this work is the analysis of the spatial variations within the fire occurrence models to explore possible local characteristics and regional patterns. For this we used GWR, which assumes non-stationary relationships between the explanatory variables and fire occurrence. Given the large territory of Spain with important climatic and socioeconomic differences – for example between the northern and the southern regions, the Atlantic and Mediterranean areas, or between the mountains, the large plains and the river depressions – we hypothesize that some explicative factors should show region-specific trends, deviating from global or national patterns. In addition, we assumed that a unique stationary model for Spain would be “notably influenced by the high fire occurrence of the Galicia region (northwest of the country). This area contains 11 % (850) of all municipalities in Spain, but 70 % (152 891) of the forest fires, and thus creates a spatial imbalance in the global model” (Martínez et al., 2009, p. 1251). These obvious premises have scarcely been tested in Spain using quantitative models. Therefore, to explore the spatial variability, we focus on the assessment (i) of variables presenting contradictory signs to the global coefficients, (ii) of areas where we observe unusually or unexpectedly high or low local coefficients, and (iii) of variables presenting a positive or negative influence in the model; finally, (iv) we try to deduce, if possible, the cause of those spatial patterns.

The third novelty of this paper with respect to our previous study (Martínez et al., 2009) was the consideration of missing explanatory environmental variables regarding climate, vegetation and topography. These structural environmental variables are essential to enable fire ignition and they are the basis on which the remaining socioeconomic, historical, land use and landscape variables interact. Besides, missing

response variables could be the cause of the unexplained spatial variations in the model parameters (Fotheringham et al., 2002; Koutsias et al., 2012). For this reason we have specifically tried to consider all key explanatory factors, including the environmental ones.

The starting point of creating fire occurrence models is to identify the most critical factors and then define and gather datasets to generate quantitative models. We identified factors based on previous literature reviews about fire causes and fire modelling (Leone et al., 2003, 2009; Martínez et al., 2004, 2009; Vélez, 2009), where the theoretical or expected relationships between fires and each factor/variable in Spain and Mediterranean countries were explained. Then, each of those explanatory factors was measured as a numeric indicator (direct or surrogate) from available datasets.

However, model building in this study is not fully “conceptual” because the final variable selection is obtained by semi-automatic statistical techniques. Besides, this paper does not intend to build a “mechanistic” or “cause–effect” model that explains human fire occurrence in the different environments of Spain. For that objective it would be better to use other approaches, for example, such as defining different environmental regions or study areas inside Spain and building specific models for each region, and subsequently comparing which are the most influential variables for each. Nor do we intend to analyze the statistical and spatial interactions between explanatory variables within the global models.

### 3 Materials and methods

#### 3.1 Study area and fire database

As the dependent variable, the numbers of human-caused fires occurring within a 25 yr period (1983–2007) were computed for each of the 7638 municipalities of the Spanish peninsula (487 000 km<sup>2</sup>) analyzed. This is another improvement over the previous study of Martínez et al. (2009) in which a 13-yr series from 1987 to 2000 was used. These data were obtained from the Spanish Forest Fire Report Database, one of the best and longest fire statistics in Europe (Leone et al., 2009). A total of 383 657 fire events has been gathered and considered in the database, regardless of their size. A binary variable (fire/no fire) for each municipality was derived to develop logistic models, and a continuous variable (fire density or the total number of fires in the period divided by the area of each municipality in km<sup>2</sup>) was estimated to build linear regression models, in this case selecting only the 6993 municipalities in which one or more fires were registered during the study period. Log transformation was applied to convert the original fire density values (Fig. 1c) to approximate a normal distribution (Fig. 1d), since the original count data would be more appropriately modelled with Poisson or Negative Binomial models depending on their variance to mean ratio (Cardille et al., 2001). The spatial distribution of both

dependent variables is shown in Fig. 1a and b, revealing critical regions for fire occurrence, especially in the NW of the country, and also along the Mediterranean coast and in some mountain ranges in the centre.

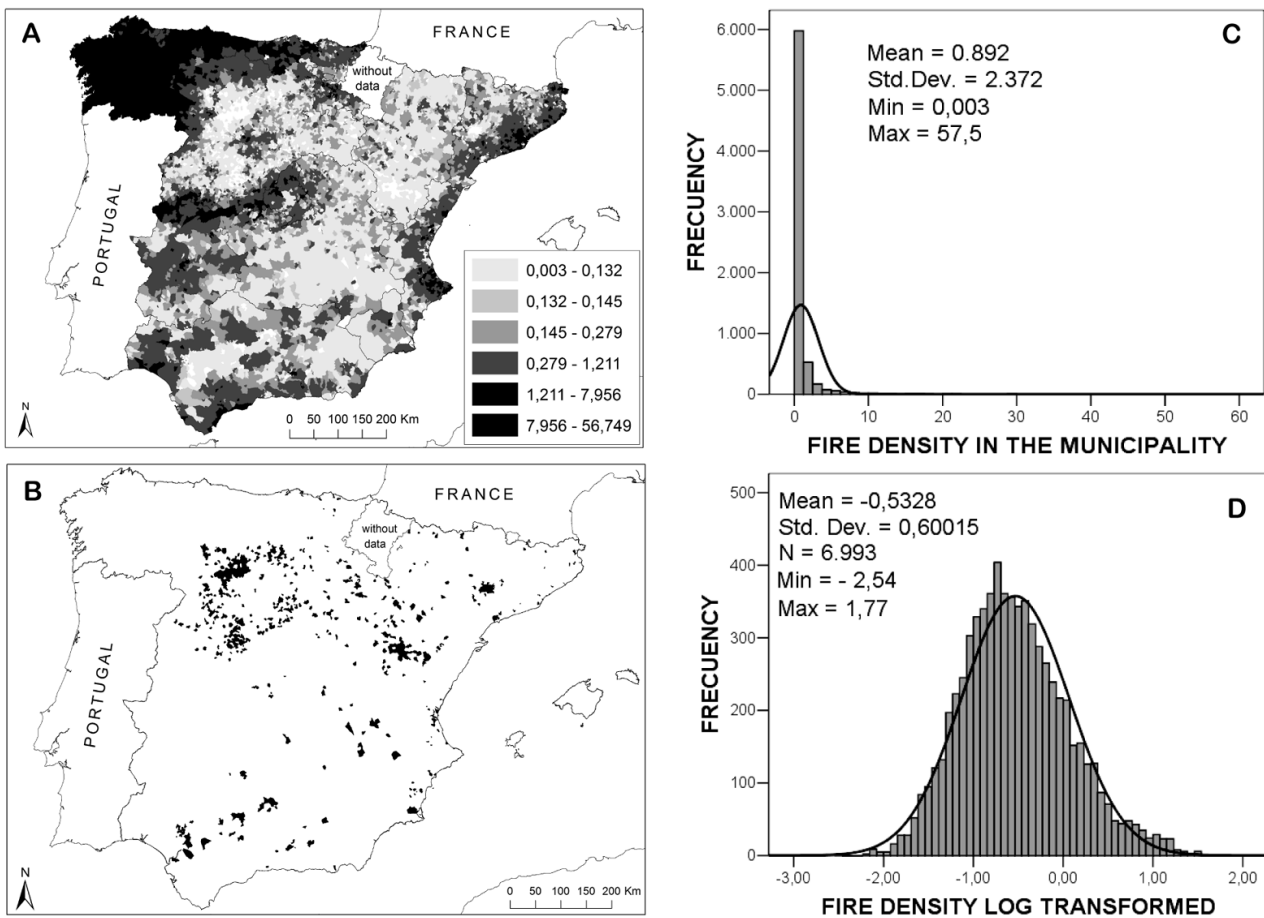
In the previous study the dependent variable was defined as the cumulative number of fires in the studied period divided by the forest area of each municipality. Instead, the flammable land cover (both vegetation and crops) are here considered as explanatory variables in order to analyze their influence and weight inside the models, and not as part of the dependent variable.

#### 3.2 Independent variables

The independent variables used in the analysis were composed initially by 29 socioeconomic and demographic indicators together with agricultural and land cover statistics compiled in Martínez et al. (2009). The identification of these variables was based on experts’ interviews, analyses of fire reports and causality statistics, and an extensive literature review. Some of the factors could not be estimated directly or from surrogate variables, while some others were not available for all the regions. Table 1 in Martínez et al. (2009) lists these variables along with their theoretical relationships with fire ignition factors and the literature source, when available. Additionally, for the present study 6 new environmental variables were added referring to topographic characteristics (mean altitude and slope), climatic indicators (summer temperature and mean annual precipitation obtained from Ninyerola et al. (2005), using the available station data set with more than 15 and/or 20 yr), and forest vegetation statistics (total wildland area and the wildland area without tree cover, both obtained from the Forest Map of Spain-MFE50 developed between 1997–2006). Total wildland area included tree-covered areas (standing forest), shrublands and grasslands and theoretically this variable is supposed to be more related to fire presence (binary model). Wildlands without tree cover only comprise shrublands and grasslands. We hypothesize that these types of areas are more strongly correlated with the fire density (linear model). All 35 variables were compiled and calculated at the municipality level for the peninsular territory of Spain, with the exception of the region of Navarre, and all of them selected after checking for multi-collinearity as described also in Martínez et al. (2009, p. 1244).

#### 3.3 Global models using classical regression

Both predictive models, based on OLS and binary logistic regression, were calculated in SPSS using automatic stepwise forward procedures for variable selection in combination with manual modification (i.e. selection using the “introduce method”). All cases were checked for potential collinearity problems of the selected variables by calculating the correlation matrix and applying other common statistical



**Fig. 1.** Map and histograms of dependent variables. Spatial distribution of the fire density in the Spanish municipalities with more than 1 event (A). The log transformation was applied to convert the original values of the dependent variable (C) to an approximate normal distribution (D). The map (B) shows the spatial distribution of the municipalities without fires used for the binary logistic modelling.

**Table 1.** Model parameters and sensitivity analysis for ordinary logistic model: ranking of influence of the input variables (the lower the ranks, the more important).

VARIABLE DESCRIPTION	MODEL PARAMETERS						RANKING CRITERIA				
	Coef. <i>B</i>	Std. Coef. <i>B</i>	Wald	Change in $-2 LL$	Exp ( <i>B</i> )	(i)	(ii)	(iii)	(iv)	(v)	Score
<b>FOR_P</b> % Forest and wildland surface	0.034	0.979	160.7	200.84	1.034	1	1	1	1	3	7
<b>DIS_50_91</b> Rural exodus: population decrease between 1950 and 1991	0.021	0.588	89.8	98.27	1.021	3	2	2	2	4	13
<b>ICFSUP_P</b> Forest/cultivated land interface area	0.099	0.506	55.0	59.71	1.104	4	4	3	3	2	16
<b>NOGES_PF</b> Forest area with less management and planning over time	0.011	0.327	58.1	55.47	1.011	6	3	4	4	7	24
<b>P_A</b> Mean annual precipitation	0.003	0.687	40.6	48.09	1.003	5	6	7	6	1	25
<b>T_SU</b> Mean summer temperature	0.158	0.386	37.4	38.58	1.171	2	5	6	5	8	26
<b>DIS_SAU</b> Decrease in agricultural area between 1989 and 1999	0.012	0.247	23.8	22.09	1.012	7	7	5	7	6	32
<b>CL21_PM</b> CORINE: agriculture but with significant areas of natural vegetation	0.0149	0.163	7.4	8.59	1.015	8	9	9	9	5	40
<b>POT_DEN</b> Population potential	-0.0002	-0.155	17.7	18.54	1.000	9	8	8	8	9	42

Notes: intersect (constant) = -3.887. Ranking criteria: (i) standardized coefficients; (ii) Wald statistic; (iii) step at which the variable was input into the model in a forward stepwise automatic procedure; (iv) change in log of likelihood ( $-2 LL$ ) when the variable was removed from the model; (v) odds ratio or the exponential of the logit coefficient *B* (Exp (*B*)).

tests such as tolerance coefficient, variance inflation factor (VIF) (Krebs et al., 2012) and eigenvalue analysis (SSTARS, 2012). The regression models were built using the standardized Z-scores for the dependent and independent variables. Additionally, the normal distribution of residuals and the lack of systematic patterns were checked for OLS. Clustering of over and/or under predictions is, for example, evidence that at least one key explanatory variable is missing. For these reasons we analyzed the histogram, the scatterplot, the normal  $Q - Q$  plot and the residual maps.

To evaluate the influence of individual variables in the models, several criteria were computed and analyzed globally: (i) a simple calculation of the standardized coefficients according to the method of Menard (2010, p. 89); (ii) the t-statistic and its level of significance, although in the case of logistic regression we used the Wald statistic; (iii) the step at which the variable was input into the model; and (iv) the change in the  $R^2$  when the variable was removed from the model (the greater the change, the more important the variable). In the case of logistic regression, we used the change in logarithm of likelihood ( $-2 LL$ ); and (v) the odds ratio or the exponential of the logit coefficient  $B$  ( $\text{Exp}(B)$ ) for the logistic model case.

### 3.4 Local models using GWR

To overcome the assumption of stationarity we applied the GWR approach using the independent variables of global regression, both for the linear and the logistic model. All analyses were implemented within GWR 3.0.1 software for Windows (Fotheringham et al., 2002; Charlton et al., 2003) using both the adaptive (nearest neighbours) and the fixed (distance) kernel types, with the minimization of the corrected Akaike Information Criterion ( $AIC_c$ ) being the criterion to determine the optimal bandwidth size of the kernel functions. This parameter ( $AIC_c$ ) was also used to compare the global OLS or logistic model with the local GWR model. As a complement, the ANOVA tests the null hypothesis that the GWR model, in the linear approach, represents no improvement over the global OLS model.

The main output from GWR for each observation point is a set of parameter estimates (local coefficients for each independent variable) and associated diagnostics (standard errors, influence index, Cook's D statistics, local  $R^2$  statistic, and local standard deviation) that can be visualized within a GIS environment (Charlton and Fotheringham, 2009). Detailed analysis of these maps allowed us to better understand and explore the spatial variability of the explanatory factors, as local  $R^2$  values show the performance of the GWR model in different areas. Additionally, GWR software includes two tests to determine whether the local parameter estimates are significantly stationary or not. Firstly, the variables might exhibit non-stationarity if the inter-quartile range (25 % and 75 % quartiles) of the GWR parameters is greater than  $\pm 1$  standard deviation (SD) of the equivalent global OLS param-

eters (Fotheringham et al., 2002; Wang et al., 2005). Secondly, significance of the spatial variability in the local parameter estimates can be examined by a Monte Carlo test, but only in the case of linear GWR, since this test is not available for logistic GWR.

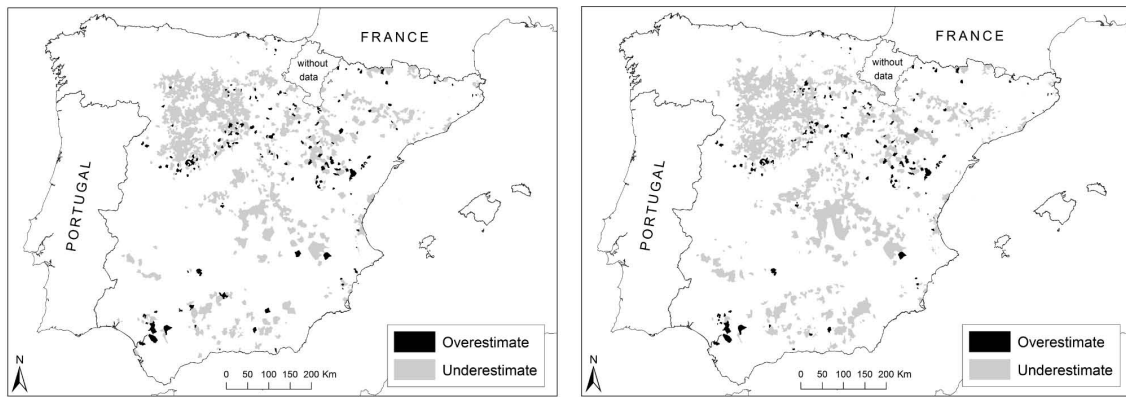
Similar to OLS regression, some spatial autocorrelation statistics for the residuals of the models have been estimated using Moran's I index of spatial autocorrelation. This made it possible to explore their spatial structure and identify whether GWR captured the spatial pattern of the residuals. If the residuals were autocorrelated then the results of the OLS regression analysis would violate one of the assumptions of OLS regression and the regression analysis would be unreliable. In the case of logistic regression, we computed the Average Nearest Neighbour Distance Index (ANND value) included in ArcGIS Desktop 10 in the Spatial Statistics toolbox. With an index  $< 1$  the pattern would tend towards clustering, while if  $> 1$  the trend is toward dispersion or competition. The interval range is from 0 to 2.14. (ArcGIS Desktop 10 Help).

## 4 Results

### 4.1 Classical regression models

After collinearity analysis we decided not to introduce the variables "slope" and "population occupied in agriculture" into the regression procedure. Instead, we introduced the variable "agricultural areas but with significant areas of natural vegetation". The stepwise procedure for the binary logistic regression selected 9 significant variables for the final model, which successfully classified 76.4 % of the total observations using the estimated optimal cut-off point of 0.91, which corresponds to the intersection of the two lines in which sensitivity and specificity are equal (Vasconcelos et al., 2001). Among the nine explanatory variables identified as critical by the analysis, the most important variables were the forest surface, population decrease and forest-cultivated land interface. Mean annual precipitation and mean summer temperature were also relevant (Table 1). The spatial distribution of the residuals (over and under estimations) of the logistic model (Fig. 2) shows that the spatial pattern of the errors is not very clear because they are dispersed through different regions of the country. However, some areas were error-free, particularly in the north, northwest and some parts of the centre and west. Most of the errors are underestimations (Table 2) because it was more probable that at least one fire had occurred during the 25-yr period than for no fires at all. The Average Nearest Neighbour Distance Index (ANND) showed the residuals tended towards clustering (0.77) and the Z-score of  $-18.6$  indicates there is less than a 1 % likelihood that this clustered pattern could be the result of a random process.

In the case of OLS regression, the model selected 23 variables as significant using an automatic stepwise procedure.



**Fig. 2.** Municipalities where prediction and observation data did not agree; either fire is predicted when it is not observed (overestimated) or fire is not predicted when it is observed (underestimated). Ordinary logistic regression on the right and GW logistic regression on the left.

**Table 2.** Residual spatial performance of logistic models using the Average Nearest Neighbour Distance Analysis.

	Ordinary Binary logistic model			GW Binary Logistic model		
	Total errors	Over-estimated	Under-estimated	Total errors	Over-estimated	Under-estimated
Number of cases	1797	158	1639	1513	170	1343
Nearest neighbour ratio	0.77	0.62	0.74	0.70	0.69	0.66
z-score	-18.60	-9.08	-20.20	-22.15	-7.62	-23.89
p-value	0.000	0.000	0.000	0.000	0.000	0.000

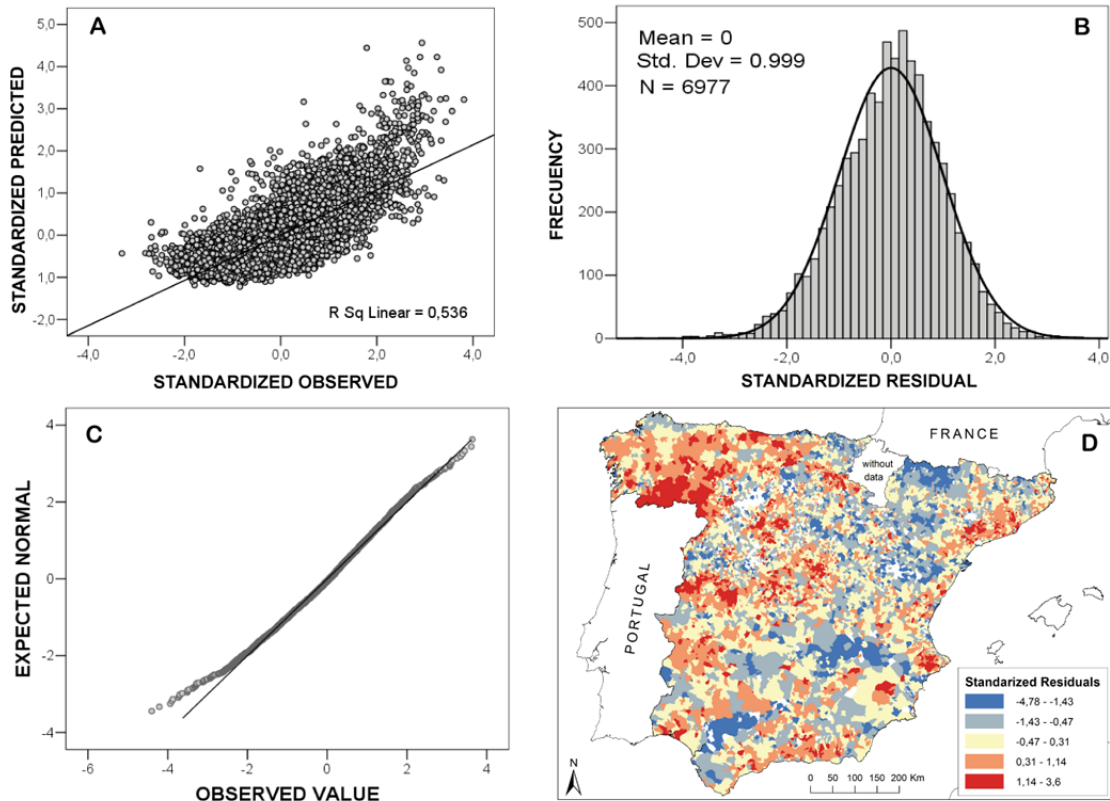
To simplify the model we selected the first nine most explanatory, plus three others in positions 11, 15 and 18 (decrease in number of owners of agrarian holdings, % owners of agrarian holdings > 55 yr, and density of agricultural machinery, respectively) that, in our opinion, included relevant aspects of agrarian structure, as reported by different regional studies. Consequently, the final model consisting of 12 variables (Table 3) explained 53 % of the variation of the dependent variable (adjusted  $R^2 = 0.53$ ). Among these variables, mean annual precipitation, density of agricultural properties, mean altitude, population decrease and non tree-covered forest surfaces were the most explanatory. For the residuals of this OLS regression model, the Kolmogorov-Smirnov test value was low (0.028) but still significant ( $p = 0.000$ ), showing that the residuals fit the normal curve poorly. However, as can be observed in the histogram (Fig. 3b), the residuals with a mean value close to 0 and a SD of 0.99 approximate acceptably well to the shape of the normal curve. The Normal  $Q - Q$  plot (Fig. 3c) represents the expected values in a straight line when the data are normally distributed. In this case, the residuals fit properly except for low observed values (low fire densities). A clustered pattern can be observed in the distribution map of the residuals (Fig. 3d) in parts of the country, although there is no clear systematic pattern. The over-predicted cases (negative values) were more concentrated in some inland areas of the eastern part of the Iberian

Peninsula. Under-prediction was more dispersed with some areas especially clustered in the NW.

Both models, logistic and OLS, are complex with a high number of variables, and for some variables the effect of introducing them in the model (measure by the change in  $R^2$  or the change of  $-2LL$ ) is very weak, as can be seen in Tables 1 and 3, although still significant. In any case, the high number of variables in the model was considered in agreement with the objective of identifying the factors that are more significant to explain fire risk, rather than obtaining parsimonious models with very few variables.

#### 4.2 Geographically weighted regression models

The GWR results showed that local models based on GWR generally fit better than global models based on classical OLS or logistic regression, while the number of effective parameters increased considerably from 10 to 29.7 in logistic, and from 13 to 63.03 or 158.34 for linear models. Based on the minimization of the corrected Akaike Information Criterion ( $AIC_c$ ), the best fixed bandwidth size in the case of logistic GWR was 219 km ( $AIC_c = 3321.19$ ). For the linear GWR, the best bandwidth size for the fixed mode was a distance of 154 km ( $AIC_c = 5898.6$ ), while for the adaptive mode it was 5724 nearest neighbours ( $AIC_c = 6395.5$ ). Models like this, with a high number of neighbours, tend to have a poor fit and present an oversmoothed pattern, as could be



**Fig. 3.** Residual analysis of the OLS regression model: scatterplots between observed and predicted observations (A), histogram data plots of the standardized residuals (B), normal  $Q - Q$  plot of the standardized residuals (C) and map of the standardized residuals (D).

**Table 3.** Model parameters and sensitivity analysis for ordinary linear regression model (OLS): ranking of influence of the input variables (the lower the ranks, the more important).

VARIABLE DESCRIPTION	MODEL PARAMETERS					RANKING CRITERIA				
	Coef. <i>B</i>	Std Coef.	<i>t</i>	Sig.	$R^2$ Change	(i)	(ii)	(iii)	(iv)	Score
<b>P.A</b> Mean annual precipitation	0.0013	0.331	33.5	0.000	0.304	1	1	1	1	4
<b>PAR_SEXP</b> Agricultural land fragmentation: density of agricultural plots	0.0020	0.185	-16.2	0.000	0.092	2	3	2	2	9
<b>ALT_MEAN</b> Mean altitude of municipality	-0.0005	-0.178	17.8	0.000	0.046	3	2	3	3	11
<b>DIS_50_91</b> Rural exodus: population decrease between 1950 and 1991	0.0057	0.161	11.5	0.000	0.035	4	6	4	4	18
<b>DESAR_P</b> % Wildlands without tree cover (mainly shrub and grasslands)	0.0077	0.147	15.6	0.000	0.025	5	4	5	5	19
<b>FRAG7 × 7</b> Landscape fragmentation index using a 7 × 7 kernel	0.1810	0.111	15.4	0.000	0.012	6	5	6	6	23
<b>ENTSIN_M</b> Density of human settlements	0.0024	0.081	7.8	0.000	0.008	7	7	7	7	28
<b>ROAD_DEN</b> Density of roads	0.0003	0.074	7.4	0.000	0.005	8	8	8	8	32
<b>IUF_DEN</b> Urban/forest interface density	0.0005	0.059	6.2	0.000	0.003	9	10	9	9	37
<b>DIS_TIT</b> Decrease in number of owners of agrarian holdings 89–99	-0.0022	-0.056	-6.2	0.000	0.002	10	9	10	10	39
<b>MAQUIN_D</b> Density of agricultural machinery	0.0088	0.034	4.2	0.00003	0.001	12	11	11	11	45
<b>TIT_55_P</b> Rural ageing: % owners of agrarian holdings > 55 yr	0.0026	0.035	3.1	0.00195	0.001	11	12	12	12	47

Note: intersect (constant) = -1.2073. Ranking criteria: (i) standardized coefficients; (ii) *t* statistic; (iii) step at which the variable was input into the model in a forward stepwise automatic procedure; (iv) change in the  $R^2$  when the variable was removed from the model.

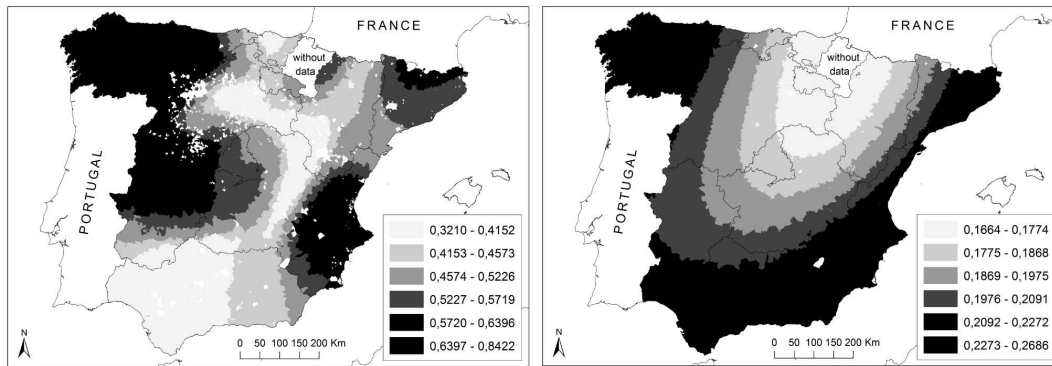


Fig. 4. Model fitting maps: local squared  $R$  for GW lineal (left) and GW logistic (right).

observed in the resulting maps. For that reason, we finally selected a kernel size of 1300 nearest neighbours, which showed a better fit ( $AIC_c = 5063.72$ ) when trying to better capture the regional variations within the country, avoiding both over- and under-smoothing. In relation to the adaptive kernel, the manually chosen value of 1300 nearest neighbours approximately represents the number of municipalities of two contiguous average regions in Spain. Regarding the fixed kernel of 154 km defined automatically by statistical criteria for GWR software, this is also considered appropriate to capture regional variations, since the mean area of the Spanish regions (autonomous communities) is 34 524 km<sup>2</sup> (e.g. Catalonia or Extremadura), which corresponds to a circle of about 145 to 160 km radius.

Comparing the fitting of the OLS and GWR models, the GWR logistic model, using a fixed bandwidth of 219 km, correctly classified 78.4 % of the observations compared to 76.4 % of the ordinary logistic regression. This improvement is not as high as expected, but it is significant as the deviance ( $-2LL$ ) improved from 3431.2 to 3261.4 and the  $AIC_c$  from 3451.2 to 3321.19. The optimal cut-off point for the classification of this GW logistic model according to the graph of sensitivity versus specificity is 0.90. For the linear approach, the explanatory power of the OLS model increased from 53 % to 67 % in the case of the adaptive mode, using a bandwidth of 1300 nearest neighbours, and 62 % in the case of the fixed mode using a bandwidth of 154 km. The adaptive mode gave slightly better results, as indicated by the coefficient of determination with a 14 % improvement, while in the case of logistic GWR it was only 2 %. The  $AIC_c$  enhanced considerably using GWR (from 7440.3 to 5063.7). The  $F$ -value of the ANOVA test suggests that the GWR model is a significant improvement on the global OLS model in Spain, at a confidence level less than 0.01 (99 %), for both fixed and adaptive models.

Figure 4 shows local  $R^2$  values for GW linear and GW logistic models indicating the areas where the predictions of the models are better. In both cases, best fits were found in the northwest and some eastern areas of the Mediterranean coast

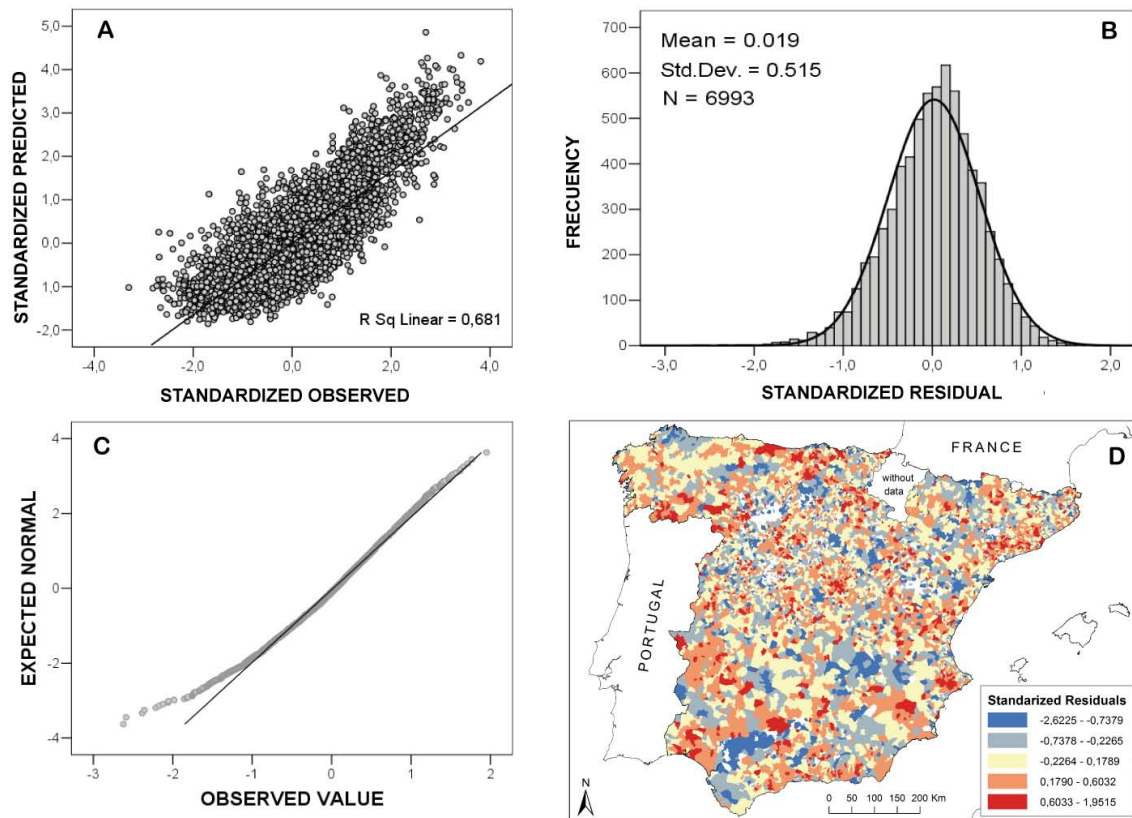
where there is usually high fire occurrence (check Fig. 1a). However, these maps are too oversmoothed in capturing local variations, especially in the logistic GWR model. The trend of the residuals of the logistic model towards clustering did not significantly decrease from the ordinary model to the GWR model according to the Average Nearest Neighbour Distance Analysis (Table 2), and there was only a minor improvement, especially in the overestimation errors. Although in the GW logistic model there were fewer errors, the spatial distribution was very similar to the ordinary logistic model (Fig. 2). Also, analysis of the linear GW regression model residuals revealed similar characteristics to the global OLS model, with a mean value of 0.01 and a SD of 0.51, acceptably following the shape of the normal curve (Fig. 5b). The Kolmogorov-Smirnov test value was low (0.03) but still significant ( $p = 0.000$ ), showing that the normal fit was poor. The residuals fitted properly except for low fire density values according to the normal  $Q - Q$  plot (Fig. 5c). However, the scatterplot was more compact along the tendency line and the standardized residual map (Fig. 5d) showed a more dispersed distribution through the study area in comparison to the OLS model (Fig. 3d), without any evident systematic pattern. These analyses indicated a slightly better performance of the GWR model.

### 4.3 Regional and local variations

The results of the Monte Carlo test on the local estimates pointed out a significant spatial variation (at 0.1 % significance level) in the local parameter estimates for all the variables of both linear GWR models (fixed and adaptive). Besides, all variables in the linear and logistic models showed evidence of spatial variability (non-stationarity) across the study area since the inter-quartile range (25 % and 75 % quartiles) of the GWR parameters was greater than  $\pm 1$  SD of the equivalent global OLS parameters.

Local coefficient estimates for each explanatory variable are presented in Fig. 6 for logistic GWR, and in Fig. 7 for adaptive linear GWR. Negative coefficients are represented by cold colours (green to blue) and positive coefficients with





**Fig. 5.** Residual analysis of the GW linear regression model: scatterplots between observed and predicted observations (A), histogram data plots of the standardized residuals (B), normal  $Q - Q$  plot of the standardized residuals (C), and map of the standardized residuals (D).

warm colours (orange to red). The objective of these maps is to explore the spatial variability and to understand better local and regional variations of the fire occurrence causal factors in Spain, developed in the discussion section.

**4.4 Spatial autocorrelation of residuals**

Spatial correlograms of the residuals of the linear models (Fig. 8) show that there is significant spatial autocorrelation of the residuals of the OLS regression model up to a distance of 600 km, while for the residuals of the GWR model the autocorrelation has been reduced significantly but still exists in relative short lag distances, up to approximately 110 km (Table 4). Less structured residuals have been observed in other studies dealing with GWR (Koutsias et al., 2010), indicating that although the method does not directly address spatial autocorrelation issues (Jetz et al., 2005), it provides a solution to the problem of spatially autocorrelated errors (Propastin and Kappas, 2008).

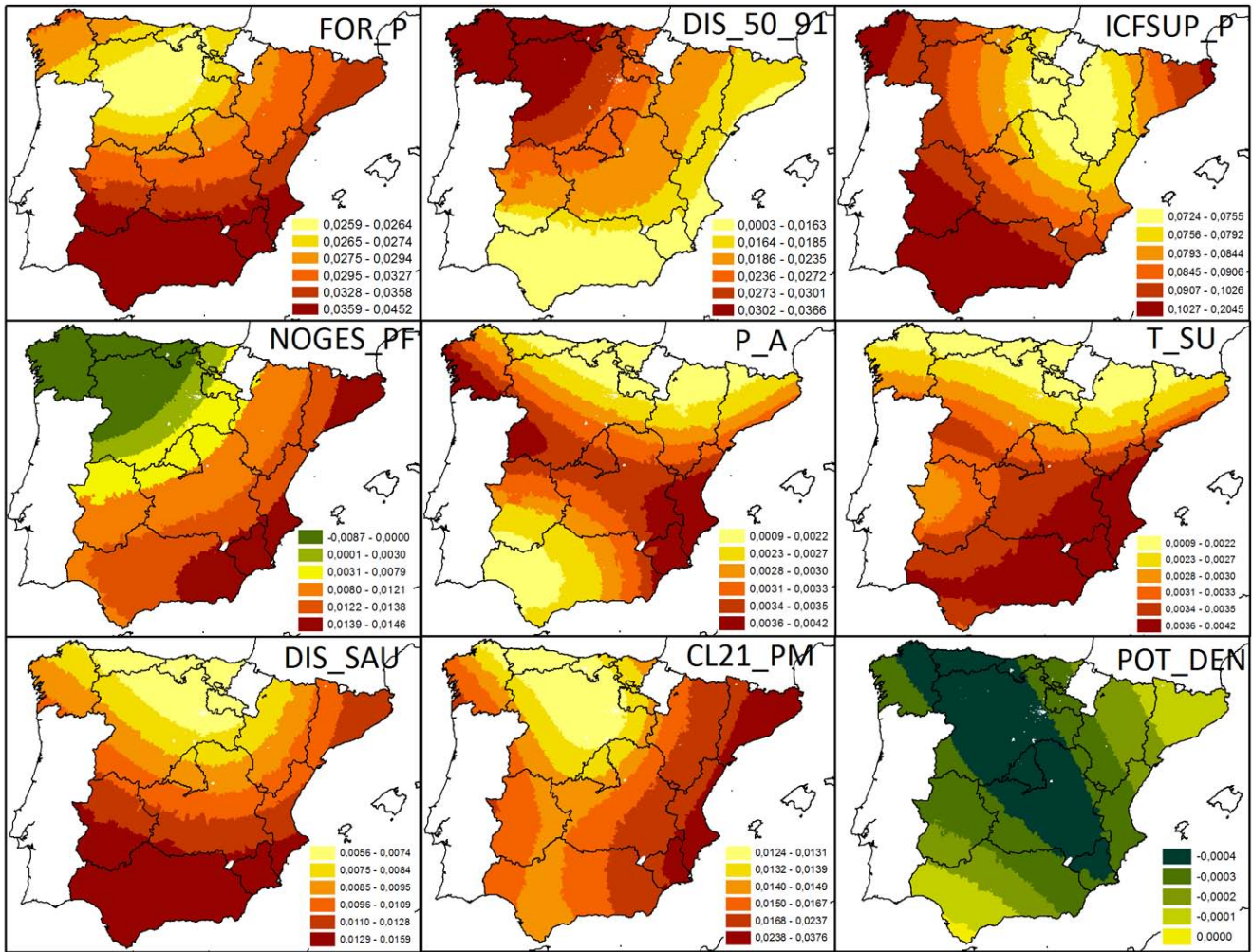
**5 Discussion**

In both regression modelling approaches there were important variables related with land and population abandonment,

agrarian activities, or development processes, in addition to forest properties and climatic variables. However, only two variables, precipitation and population decrease, were common between the two approaches, indicating different underlying mechanisms for fire presence and for fire density at the community level. In this discussion we analyze the most important explanatory variables for each model and explore their spatial variations according to GWR local parameters (Figs. 6 and 7). Some variables presented high variability in explaining the dependent variable, occasionally even being contradictory to the global coefficients.

**5.1 Driving factors of long-term fire presence**

The percentage of wildland area was the most important factor to discriminate non fire-prone from fire-prone municipalities (defined as those in which at least one fire was observed during the 25-yr period studied). This is reasonable since the probability of fire ignition and spread was very low in places with a very low percentage of forest and natural cover, as fuels are very scarce or non-existing. The influence of this variable was higher in the south of the country, as observed in Fig. 6. Another important variable was forest-cultivated land interface (ICFSUP\_P), which is related to agricultural activities where fire is frequently used in arable and crop lands,



**Fig. 6.** Local coefficients for GW binary logistic model using a fixed bandwidth of 219 km. Negative coefficients are mapped with cold colours (green) and positive with warm colours (orange to red). Variable names and their descriptions are in Table 1.

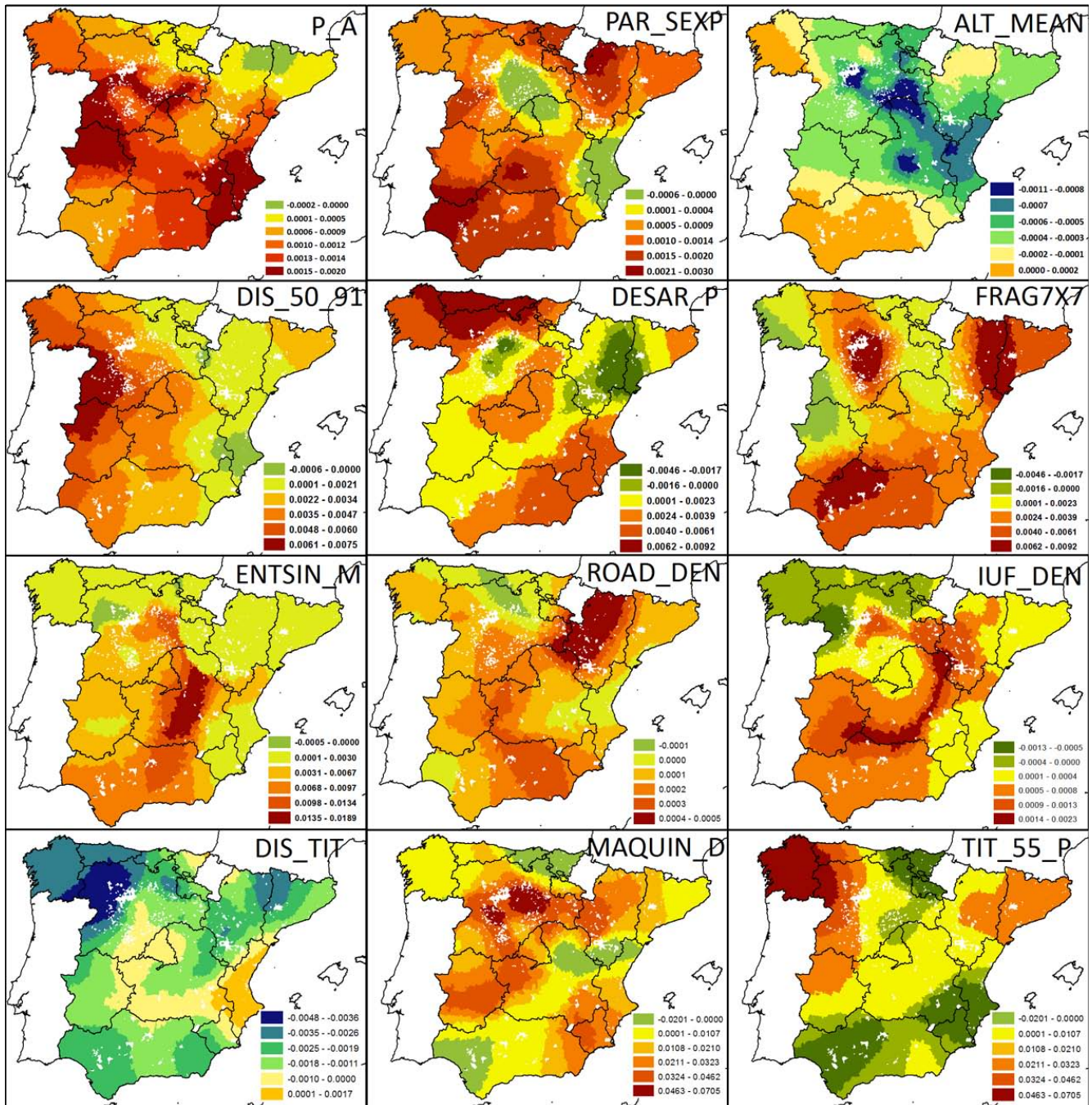
**Table 4.** Moran’s Index Summary for different band distances on the Linear GWR residuals.

Threshold Distance	90 km	100 km	105 km	110 km	120 km	154 km
Moran’s Index	0.007316	0.003598	0.001945	0.00088	−0.00041	−0.00021
Expected Index	−0.000143	−0.000143	−0.000143	−0.00014	−0.000143	−0.00014
z-score	8.939	4.95545	2.898472	1.48896	−0.421	−0.125
p-value	0.000	0.000001	0.00375	<b>0.13650</b>	<b>0.674</b>	<b>0.900</b>
residual pattern interpretation	clustered	clustered	clustered	random	random	random

very close to or intermixed with forest areas. This variable was less relevant in the Iberian Mountain System and Ebro River Depression. The importance of the forest–agriculture interface (FAI) in forest fires in Spain was observed in previous studies (Martínez et al., 2009, Ortega et al., 2012; Gonzalez-Olabarría et al., 2012), which found that the landscapes most vulnerable to fire were those with fine-grained forest–agriculture mixtures or mosaics, where the human-

caused fires were more intense than homogeneous and non-fragmented landscapes.

Variables DIS\_50\_91 (population decrease between 1950 and 1991) and DIS\_SAU (decrease in agricultural area between 1989 and 1999) were positively correlated with the occurrence of at least one fire event. Both variables can be associated with abandonment of land and traditional activities and the movement of population from rural and mountainous areas to lowlands and urban areas. A consequence of



**Fig. 7.** Local coefficients for adaptive GWR linear model using a bandwidth of 1300 nearest neighbours. Negative coefficients are mapped with cold colours (green to blue) and positive with warm colours (orange to red). Variable names and their descriptions are in Table 3.

land abandonment is fuel build-up. Instead, according to the positive correlations observed for these variables, in municipalities with population reduction and land abandonment, fires were expected in cases where the decrease is lower. Under this demographic and social context, areas maintaining a relatively higher agricultural population are more fire prone. This is an example of the contradictory types of rela-

tionships between the explanatory and response variables in wildfire occurrence modelling. According to the local coefficient maps in the NW, the occurrence of at least one fire is more closely associated to the population presence than the rural exodus or land abandonment (DIS\_50\_91), while in the south, the presence of agricultural land (DIS\_SAU) is more influential (Fig. 6).

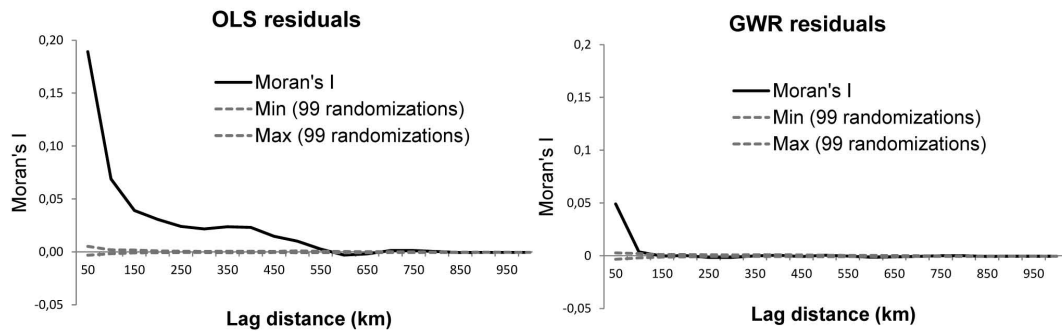


Fig. 8. Spatial correlograms of the residuals of the OLS linear regression (left) and GWR regression modelling (right).

Population and agricultural area decrease are also closely related with the population potential (POT\_DEN), a similar concept of population density or human presence, which is further associated with the probability of fire ignition and area burned. This has a positive influence in many studies (Cardille et al., 2001; Maingi and Henry, 2007; Romero-Calcerrada, 2008; Catry et al., 2009; Sebastian-Lopez et al., 2008; Martínez et al., 2009; Marques et al., 2011; Nunes, 2012) or a negative relationship for some areas in other studies (Narayananaraj and Wimberly, 2012; Sá et al., 2011). Additionally, the previously mentioned variables were also related with the CORINE land use class “agriculture but with significant areas of natural vegetation” (CL\_21\_PM), showing that fire occurrence was more likely in municipalities where agricultural and forest areas are intermixed, similar to what has been reported by Ortega et al. (2012). Recently, when trying to explain the extreme 2007 fires in the Greek Peloponnese, Koutsias et al. (2012) observed that the CORINE land cover category “agricultural land, highly interspersed with significant areas of natural vegetation” was the most affected by fire, reflecting the encroachment of natural vegetation in abandoned fields and also recent patterns of evolution in the wildland–rural interface where agricultural land is increasingly intermixed with natural vegetation.

Together with land abandonment and population decrease, the economic value of lands and forests was identified as a factor of human-caused fires due to a decreasing involvement in conservation and land management by the remaining rural population. In this sense, the NOGES\_PF variable in the model was positively correlated with fire occurrence. This variable measures the percentage of forest surface with less management, control and planning over time, which in Spain is the private forest land, land belonging to local authorities with free use, consortiums and neighbouring forests. All these kinds of properties have a generally worse conservation and protection status than national, regional or public forest. Local coefficients for this variable were positive in the Mediterranean coast and negative in the NW. Padilla and Vega-García (2011) found that several variables related to forest ownership (private, public and communal areas) were significant for the northern ecoregions of Spain.

Finally, climatic variables were also found to be relevant factors to explain fire occurrence. Mean summer temperature and mean annual precipitation are important factors, especially in the warmer areas of the E and SE. Many studies (Shyphard et al., 2008; Drever et al., 2008; Vilar et al., 2010; Padilla and Vega-García et al., 2011; Oliveira et al., 2012, Sá et al., 2011; Narayananaraj and Wimberly, 2012; Nunes, 2012) selected several climatic variables as very significant in their fire models – some related to precipitation, such as fire-season and off-season precipitation, precipitation seasonality, soil water storage and soil moisture anomaly – and others related to temperature, especially the maximum temperature in the driest season.

## 5.2 Driving factors of long-term fire density

Summer temperature was not a significant factor to explain fire density in linear regression, unlike in the logistic model (fire/no fire incidence). However, the mean annual precipitation was the most important factor to explain forest fire density (Table 3). Local coefficients for this variable were positive across almost the entire country, especially in the SE and some parts of the inland west, which may be related to the impact of rainfall on fuel availability, particularly in the dry SE regions of Spain. The exception was observed in the NE and central Pyrenees (negative local coefficients) where rainfall occurs also in the summer and therefore the fire season tends to be shorter, although this also happens in other parts of the country where positive coefficients are found. Oliveira et al. (2012) pointed out that the most important variables related with fire density distribution in the EUMed region were off-season precipitation (positive influence related to vegetation growth and fuel accumulation) and fire season precipitation, with a negative relationship limiting fire ignition and spread. Sá et al. (2011) indicate that in the drier areas of sub-Saharan Africa there is a positive relationship between fire incidence and soil water, which is important for vegetation growth.

The density of agricultural properties (PAR\_SEXP) was positively related to fire occurrence, suggesting that highly partitioned agricultural properties increased the

human-caused ignition risk. In combination with the variable “density of agricultural machinery” (MAQUIN\_D), this indicates that, the higher the number of properties and machines, the more likely conflicts and negligence become. Fire is one of the preferred tools to eliminate stubble, weeds, field margins, hedges and shrubs, and to reclaim abandoned lands (Leone et al., 2003), especially in areas where agricultural parcel density is very high and irregularly distributed in space. In Spain, more than 20% of the fires that occurred within the 25-yr period (1980–2004) were caused by intentional or negligent agricultural burnings and other burnings of shrublands to regrow or maintain pastures for livestock, although the importance of these causes could be far greater, and actually estimated at 45% (Leone et al., 2009; Koutsias et al., 2010). In addition, as explained in Martínez et al. (2009, p. 1248), in many cases mechanization implies a willingness to obtain more space and land for cultivation, and fires are one of the tools to achieve it. Also, more intensive agricultural activity, promoted by mechanization over time, may increase the need to burn more stubble, agricultural residues and prunings, as well as a higher number of ignitions produced by accidental sparks deriving from engine operation. Similar agriculture related variables have also been used in other fire modelling studies (Sebastian-Lopez et al., 2008; Catry et al., 2009). According to the spatial distribution of the local coefficients, we identified areas where the expected direction in relation to high fire density exhibited opposite trends to the global model. This is the case of variable PAR\_SEXP in the Valencia Region and the northern plateau (especially in the “Ribera del Duero” region), where local negative coefficients were found. Both areas have high fragmentation of small-holdings both in irrigated and in rain-fed arable land agriculture, but they have few forested areas and a landscape with less wildland–agrarian mosaics. However these areas present high and positive coefficients for variable agricultural machinery. Instead, in the eastern Cantabrian regions (Basque Country and Cantabria) and the Guadalquivir depression in the SW, variables MAQUIN\_D (agricultural machinery) and TIT\_55\_P (percentage of old owners of agrarian holdings) showed the opposite trends. In some of these humid Atlantic environments of the north, livestock and forestry are more important than agriculture. The southwestern Guadalquivir area presented one of the lowest indices in the number of agricultural machines compared with other irrigated areas of the country.

The mean municipality altitude (ALT\_MEAN) variable was the third most explanatory in the model. The global coefficient and most of the local coefficients throughout the country were negative, so at lower altitudes more fire densities were expected, especially in the central part of the country and the eastern coast. However, we observed a positive influence across the entire S of the country and in the NW (Galicia). Coefficients were neutral (close to 0) in the N of Aragon and the central Pyrenees in the upper mountains. Other studies showed that elevation presents contrasting relationships

with fire occurrence. Some studies found a positive influence (Catry et al., 2009; Marques et al., 2011) as a consequence of pastoral burns (renovation of pastures for livestock) or a higher frequency of lightning in higher altitudes (Vazquez and Moreno, 1998; Narayanaraj and Wimberly, 2012), while other studies found a negative correlation (Vasconcelos et al., 2001; Sebastian-Lopez et al., 2008; Gonzalez-Olabarría et al., 2012; Vilar et al., 2010; Padilla and Vega-Garcia, 2011), suggesting that lower elevations tend to be the more xeric places, with dryer fuels and less productivity. However, fuel dry-out is probably a function of the temporal distribution of precipitation, which in the Mediterranean area is very high in summer due to the seasonal drought. In addition, when altitude increases, the vegetation loading tends to decrease with more unburnable areas appearing (rocks, sparse vegetation, ice, etc.) although only over a certain height. Unlike lightning-caused fires, Narayanaraj and Wimberly (2012) detected a negative association between elevation and slope and human-caused fires in a mountain area of Washington State. Similarly, Vilar et al. (2010) found a less intense land use at high elevations in the Madrid Region. In some regions of Spain, as in other parts of the world, population, roads and some land uses responsible for the higher number of ignitions are concentrated in coastal areas, decreasing with increasing elevation (Badia-Perpinya and Pallares-Barbera, 2006). The same conclusions about how topography reflects the locations of human activities in relation to fire ignitions are indicated for a region of China by Xu et al. (2006), showing that the anthropogenic factors are closely related to fires when altitudes of forests are lower than 900 m. However, at higher elevations their influence is much lower. In other studies the topography effect has been related with fires, using variables related with roughness or terrain shape index (Dickson et al., 2006; Nunes, 2012; Padilla and Vega-Garcia, 2011; Narayanaraj and Wimberly, 2012).

As in the logistic model, population decrease between 1950 and 1991 (DIS\_50\_91) was found to be a relevant explanatory variable. Besides, the inclusion in the model of DIS\_TIT (decrease of the number of owners of agrarian holdings 89–99) reinforces the idea of the relationship between land abandonment and rural exodus, with a high fire risk (Hill et al., 2008; Nunes, 2012). The local coefficient maps of these two variables portrayed two patterns: (1) on one hand, in the east of the country, fires are related to population abandonment and rural exodus, with the resulting accumulation of fuels, but at the same time there is some maintenance of agricultural activities related to fires because the variable DIS\_TIT correlates positively with fire density; and (2), in contrast to this trend, the western region, particularly in the NW, DIS\_50\_91 has positive correlation with fires, indicating a further influence of the population presence on fires. However, in these western areas, DIS\_TIT correlates negatively to fires, so if agricultural land owners decline and agricultural activities are abandoned, fires tend to be more frequent, mainly because of greater fuel accumulation. In addition, in

this NW area (Galicia), which is the most fire affected region, we found strong positive coefficients with the variable ageing agricultural population (TIT\_55\_P), because in this region the older population is the more accustomed to use fire in farming works, as they did in their youth (Vélez, 2009). This variable also suggests the impact of the land abandonment process in increasing fire frequency. A similar trend was also observed by Nunes (2012) in Portugal, where the ageing index correlates negatively with the density of the population and is positively associated with agricultural land abandonment. The same process has been pointed out for the Mediterranean European Basin using the difference of the youth index between 1990 and 1960 (Koutsias et al., 2010) as a proxy.

Two of the variables relate fires to forestry and landscape features. Thus, more fires were found in landscapes with a large percentage of shrublands and grasslands (DESAR\_P), especially in the Cantabrian Mountains and N coast where there are numerous pastoral fires to create, maintain, or re-grow pastures for livestock (Moreira et al., 2011). Many studies have confirmed that shrubland is one of the most fire affected land cover types (Nunes et al., 2005; Sebastian-López et al., 2008; Catry et al., 2009; Moreira et al., 2009; Bajocco and Ricotta, 2008; Nunes, 2012; Oliveira et al., 2012) due to a combination of factors: “a higher rate of fire spread, a larger frequency of ignitions (e.g. to create pastures) and a lower fire fighting priority” (Marques et al., 2011, p. 783). In sub-Saharan Africa the herbaceous vegetation proportion is the variable best related with fire incidence (Sá et al., 2012). On the other hand, more fire density was found in Spanish landscapes with high fragmentation (FRAG7 × 7), especially in the three main river depressions where agriculture dominates (Duero, Ebro and Guadalquivir) and where there is less forest cover. Heterogeneous and interspersed patterns composed by spatially separated patches with different land uses presents higher ignition frequencies (Ortega et al., 2012; Ruiz-Mirazo et al., 2012).

Finally, two variables were related, in general, to human presence and accessibility. Road density (ROAD\_DEN), the same as road distance, has very often been found to be related to human accidental or negligent fires (Cardille et al., 2001; Vasconcellos et al., 2001; Badia-Perpinya and Pallares-Barbera, 2006; Yang et al., 2007; Romero-Calcerrada et al., 2008; Martínez et al., 2009; Catry et al., 2009; Vilar et al., 2010; Padilla and Vega-García, 2011; Gonzalez-Olabarria, 2012; Oliveira et al., 2012; Narayanaraj and Wimberly, 2012), as well the density and distance to human settlement (ENTSIN\_M). Specifically, the risk appeared to be higher in the urban–forest interface (IUF\_DEN) zone (Badia-Perpinya and Pallares-Barbera, 2006; Syphard et al., 2007; Catry et al., 2009; Martínez et al., 2009; Romero-Calcerrada et al., 2010; Vilar et al., 2010; Gonzalez-Olabarria et al., 2012; Narayanaraj and Wimberly, 2012) where population and human infrastructure facilities are in contact, close to or disseminated throughout the forested zones, especially in the vicinity of large cities and tourist resorts (Viegas et al., 2003).

According to the spatial distribution of the local coefficients of these variables, it might be surprising to find high coefficients in areas where the values of these variables are low, such as in the southern part of Central Spain and in the inland mountains of the south, where a low population density, few population centres and scarce wildland–urban interfaces are found. However, these few places with higher densities of human activities seem to tend to bias the model and therefore seem to be decisive for the fire occurrence in those areas.

## 6 Conclusions

In this study we built two complementary models which cover the whole range of the human-caused fire occurrence in Spain during a 25 yr period. The first model tries to predict and explain fire densities, and the second fire presence/absence. The most influential variables for both models are related to agrarian activities, land abandonment, rural exodus and development processes. Additionally, specific traits of vegetation, climatology and topography have also been very important, since they affect the initial conditions enabling fire incidence. The inclusion of these environmental variables results in an improvement over the previous model (Martínez et al., 2009), on which this study is based.

Relevant differences between both models are found because only two explanatory variables are common: mean annual precipitation and population decrease. Potentially flammable land cover types (total wildland area and agricultural/forest interfaces and mosaics) and the mean summer temperature are the main specific variables for the fire presence model. Instead, agricultural fragmentation, elevation, shrublands and grasslands, along with human structures (roads, settlements, etc.) and other rural indicators are specific variables for the fire density model.

However, these stationary models and global regression approaches seem to be insufficient to appropriately explain the underlying fire factors, because all variables selected showed significant spatial variations at the regional or local scale according to the GWR model. Nevertheless, only some of them present, in fact, very high variability or contradictory relationships with the response variable and/or the global trends. For example, the density and fragmentation of agricultural plots has a negative relationship with fires in regions with low forest areas and less wildland–agrarian mosaics, as along the E coast in the Valencia Region and the eastern part of the northern plateau (Ribera de Duero), and both areas are characterized by having small-holdings of irrigated agriculture and also rain-fed arable lands. Also, precipitation, decrease in owners of agrarian holdings, population entities or the urban–forest interface present unexpectedly high regression coefficients in areas where those variables have low original values. Thus, although precipitation seems to be a very important factor to model fire densities in the driest areas of the country, it is not that relevant in other areas

with more rainfall availability. Similarly, the percentage of forest and wildland area has a higher influence in the S of the country, which is drier and with less vegetation as compared to the N. Another interesting pattern is observed between the E and the W–NW where population presence seems to have a further influence on fires, although at the same time important land abandonment processes are observed. In the E, instead, fires seem directly more related to population abandonment and rural exodus, but also to agricultural activities, though to a lesser degree. Finally, lower altitude seems more related with the fire density along the eastern coast and in the central part of the country, unlike the pattern observed in the S and NW where higher altitudes present more fire risk. In the upper mountains of the Central Pyrenees this relation is neutral.

This analysis is another contribution to the field of fire management and fire risk assessment in the Mediterranean countries, which quantitatively and spatially demonstrated the importance of considering regional variations and local modelling as a complement to global and stationary models in order to better understand the fire problem over large study areas.

*Acknowledgements.* The authors are grateful to Ministry of Agriculture, Food and Environment of Spain for providing access to fire data base. Comments, suggestions and corrections of the editor and reviewers during the evaluation process are greatly appreciated.

Edited by: K. Chang

Reviewed by: O. Price and one anonymous referee

## References

- Avila-Flores, D., Pompa-Garcia, M., Antonio-Nemiga, X., Rodriguez-Trejo, D., Vargas-Perez, E., and Santillan-Perez, J.: Driving factors for forest fire occurrence in Durango State of Mexico: A geospatial perspective, *Chinese Geogr. Sci.*, 20, 491–497, 2010.
- Badia-Perpinya, A. and Pallares-Barbera, M.: Spatial distribution of ignitions in Mediterranean periurban and rural areas: the case of Catalonia, *Int. J. Wildland Fire*, 15, 187–196, 2006.
- Bajocco, S. and Ricotta, C.: Evidence of selective burning in Sardinia (Italy): which land cover classes do wildfires prefer?, *Landscape Ecol.*, 23, 241–248, 2008.
- Brunsdon, C., Fotheringham, A. S., and Charlton, M.: Geographically weighted regression: a method for exploring spatial nonstationarity, *Geogr. Anal.*, 28, 281–298, 1996.
- Carmo, M., Moreira, F., Casimiro, P., and Vaz, P.: Land use and topography influence on wildfire occurrence in northern Portugal, *Landscape Urban Plan.*, 100, 169–176, 2011.
- Catry, F. X., Rego, F. C., Bação, F. L., and Moreira, F.: Modeling and mapping wildfire ignition risk in Portugal, *Int. J. Wildland Fire*, 18, 921–931, doi:10.1071/WF07123, 2009.
- Cardille, J. A., Ventura, S. J., and Turner, M. G.: Environmental and social factors influencing wildfires in the Upper Midwest, United States, *Ecol. Appl.*, 11, 111–127, 2001.
- Charlton, M. and Fotheringham, A. S.: Geographically Weighted Regression. A Tutorial on using GWR in ArcGIS 9.3, available at: [http://ncg.nuim.ie/ncg/gwr/GWR\\_Tutorial.pdf](http://ncg.nuim.ie/ncg/gwr/GWR_Tutorial.pdf), National Centre for Geocomputation, National University of Ireland Maynooth, Maynooth, County Kildare, Ireland, 25 pp., 2009.
- Charlton, M., Fotheringham, A. S., and Brunsdon, C.: Geographically Weighted Regression. VERSION 2.x, User's Manual and Installation Guide. National Centre for Geocomputation, National University of Ireland Maynooth, Maynooth, 33 pp., 2003.
- Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M. P., Vilar, L., Martínez, F. J., Martín, S., Ibarra, P., De la Riva, J., Baeza, J., Rodríguez, F., Molina, J., Herrera, M. A., and Zamora, R.: Development of a framework for fire risk assessment using remote sensing and geographic information system technologies, *Ecol. Model.*, 221, 46–58, 2010.
- Dickson, B. G., Prather, J. W., Xu, Y., Hampton, H. M., Aumack, E. N., and Sisk T. D.: Mapping the probability of large fire occurrence in northern Arizona, USA, *Landscape Ecol.*, 21, 747–761, 2006.
- Drever, C. R., Drever, M. C., Messier, C., Bergeron, Y., and Flannigan, M.: Fire and the relative roles of weather, climate and landscape characteristics in the Great Lakes-St. Lawrence forest of Canada, *J. Veg. Sci.*, 19, 57–66, 2008.
- FAO: Fire Management – Global Assessment 2006, A Thematic Study Prepared in the Framework of the Global Forest Resources Assessment 2005, FAO, Rome, 2007.
- Fotheringham, A. S., Charlton, M., and Brunsdon, C.: The Geography of Parameter Space: An Investigation of Spatial Nonstationarity, *Int. J. Geogr. Inf. Syst.*, 10, 605–627, 1996.
- Fotheringham, A. S., Charlton, M., and Brunsdon, C.: Two techniques for exploring nonstationarity in geographical data, *Geographical Systems*, 4, 59–82, 1997.
- Fotheringham, A. S., Brunsdon, C., and Charlton, M.: Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, John Wiley and Sons, New York, 269 pp., 2002.
- González-Olabarria, J. R., Mola, B., Pukkala, T. and Palahí, M.: Using multi-scale spatial analysis to assess fire ignition density in Catalonia, Spain, *Ann. For. Sci.*, 68, 861–871, 2011.
- Hill, J., Stellmes, M., Udelhoven, T., Röder, A., and Sommer, S.: Mediterranean desertification and land degradation: Mapping related land use change syndromes based on satellite observations, *Global Planet. Change*, 64, 146–157, 2008.
- Hope, A. C. A.: A simplified Monte Carlo Significance Test Procedure, *J. R. Stat. Soc. Ser. B*, 30, 582–598, 1968.
- Jetz, W., Rahbek, C., and Lichstein, J. W.: Local and Global Approaches to Spatial Data Analysis in Ecology, *Global Ecol. Biogeogr.*, 14, 97–98, 2005.
- Koutsias, N., Martínez, J., Chuvieco, E., and Allgöwer, B.: Modeling Wildland Fire Occurrence in Southern Europe by a Geographically Weighted Regression Approach, in: Proceedings of the 5th International Workshop on Remote Sensing and GIS Applications to Forest Fire Management: Fire Effects Assessment, Universidad de Zaragoza, Spain, 57–60, ISBN:84-96214-52-4, 2005.
- Koutsias, N., Martínez-Fernández, J., and Allgöwer, B.: Do factors causing wildfires vary in space? Evidence from geographically weighted regression, *GISci. Remote Sens.*, 47, 1548–1603, 2010.
- Koutsias, N., Arianoutsou, M., Kallimanis, A. S., Mallinis, G., Halley, J. M., and Dimopoulos, P.: Where did the fires burn in Pel-

- ponnisos, Greece the summer of 2007? Evidence for a synergy of fuel and weather, *Agr. For. Meteorol.*, 156, 41–53, 2012.
- Krebs, P., Koutsias, N., and Conedera, M.: Modelling the eco-cultural niche of giant chestnut trees in southern Switzerland: new insights into landscape history through distribution analysis of a heritage, *J. Hist. Geogr.*, 38, 372–386, doi:10.1016/j.jhg.2012.01.018, 2012.
- Kwak, H., Lee, W., Saborowski, J., Lee, S., Won, M., Koo, K., Lee, M., and Kim, S.: Estimating the spatial pattern of human-caused forest fires using a generalized linear mixed model with spatial autocorrelation in South Korea, *Int. J. Geogr. Inf. Sci.*, 26, 1589–1602, doi:10.1080/13658816.2011.642799, 2012.
- Leone, V., Koutsias, N., Martínez, J., Vega-García, C., Allgöwer, B., and Lovreglio, R.: The human factor in fire danger assessment, edited by: Chuvieco, E., in: *Wildland Fire Danger, Estimation and Mapping, The Role of Remote Sensing Data*, Ser. Rem. Sens., 4, World Scientific Publishing Co. Pte. Ltd, 143–196, ISBN:981-238-569-X, 2003.
- Leone, V., Lovreglio, R., Martín, M. P., Martínez, J., and Vilar, L.: Human factors of fire occurrence in the Mediterranean, edited by: Chuvieco, E., in: *Earth Observation of Wildland Fires in Mediterranean Ecosystems*, Springer-Verlag, Berlin Heidelberg, 149–170, ISBN:978-3-642-01753-7, doi:10.1007/978-3-642-01754-4\_11, 2009.
- Maingi, J. K. and Henry, M. C.: Factors influencing wildlife occurrence and distribution in eastern Kentucky, USA, *Int. J. Wildland Fire*, 16, 23–33, 2007.
- Martínez, J., Martínez Vega, J., and Martín, M. P.: El factor humano en los incendios forestales: análisis de factores socio-económicos relacionados con la incidencia de incendios forestales en España, edited by: Chuvieco, E. and Martín, M. P., in: *Nuevas Tecnologías para la Estimación del Riesgo de Incendios Forestales*, Madrid, CSIC, Spain, 101–142, 2004.
- Martínez, J., Vega-García, C., and Chuvieco, E.: Human-caused wildfire risk rating for prevention planning in Spain, *J. Environ. Manage.*, 90, 1241–1252, doi:10.1016/j.jenvman.2008.07.005, 2009.
- Marques, S., Borges, J. G., Garcia-Gonzalo, J., Moreira, F., Carreira, J. M. B., Oliveira, M. M., Cantarina, A., Botequim, B., and Pereira, J. M. C.: Characterization of wildfires in Portugal, *Eur. J. For. Res.*, 130, 775–784, 2011.
- Menard, S. W.: *Logistic regression: From introductory to advanced concepts and applications*, Sage, London, 377 pp., 2010.
- Moreira, F., Vaz, P., Catry, F., and Silva, J. S.: Regional variations in wildfire preference for land cover types in Portugal: implications for landscape management to minimise fire hazard, *Int. J. Wildland Fire*, 18, 563–574, 2009.
- Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, E., Barbat, A., Corona, P., Vaz, P., Xanthopoulos, G., Mouillot, F., and Bilgili, E.: Landscape – wildfire interactions in Southern Europe: implications for landscape management, *J. Environ. Manage.*, 92, 2389–2402, 2011.
- Narayanaraj, G. and Wimberly, M. C.: Influences of forest roads on the spatial patterns of human- and lightning-caused wildfire ignitions, *Appl. Geogr.*, 32, 878–888, 2012.
- Ninyerola, M., Pons, X., and Roure, J. M.: *Atlas Climático de la Península Ibérica, Metodología y Aplicaciones en Bioclimatología y Geobotánica*. Universidad Autónoma de Barcelona, Bellaterra, Spain. 45 pp., ISBN:932860-8-7, 2005.
- Nunes, A. N.: Regional variability and driving forces behind forest fires in Portugal an overview of the last three decades (1980–2009), *Appl. Geogr.*, 34, 576–586, 2012.
- Nunes, M. C. S., Vasconcelos, M. J., Pereiram, J. M. C., Dasgupta, N., Alldredge, R. J., and Rego, F. C.: Land cover type and fire in Portugal: do fires burn land cover selectively?, *Landscape Ecol.*, 20, 661–673, doi:10.1007/s10980-005-0070-8, 2005.
- Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., and Pereira, J. M. C.: Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest, *Forest Ecol. Manage.*, 275, 117–129, 2012.
- Ortega, M., Saura, S., González-Ávila, S., Gómez-Sanz, V., and Elena-Rosselló, R.: Landscape vulnerability to wildfires at the forest-agriculture interface: half-century patterns in Spain assessed through the SISPAES monitoring framework, *Agroforest. Syst.*, 85, 331–349, doi:10.1007/s10457-011-9423-2, 2012.
- Padilla, M. and Vega-García, C.: On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain, *Int. J. Wildland Fire*, 20, 46–58, 2011.
- Poudyal, N. C., Johnson-Gaither, C., Goodrick, S., Bowker, J. M., and Gan, J.: Locating spatial variation in the association between wildland fire risk and social vulnerability across six southern States, *Environ. Manage.*, 49, 623–635, doi:10.1007/s00267-011-9796-z, 2012.
- Propastin, P. A. and Kappas, M.: Reducing Uncertainty in Modeling the NDVI-Precipitation Relationship: A Comparative Study Using Global and Local Regression Techniques, *GISci. Remote Sens.*, 45, 47–67, 2008.
- Rodríguez, J. and De la Riva, J.: Modelado de la variación espacial de los factores explicativos de la causalidad humana en incendios forestales mediante Regresión Logística Ponderada Geográficamente, edited by: Martínez-Vega, J. and Martín, M. P., in: *Tecnologías de la Información Geográfica en el contexto del Cambio Global. Libro de Actas del XV Congreso Nacional de Tecnologías de la Información Geográfica*, Madrid, Spain, 395–406, 2012.
- Romero-Calcerrada, R., Novillo, C. J., Millington, J. D. A., and Gomez-Jimenez, I.: GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid, Central Spain, *Landscape Ecol.*, 23, 341–354, 2008.
- Romero-Calcerrada, R., Barrio-Parra, F., Millington, J. D. A., and Novillo, C. J.: Spatial modelling of socioeconomic data to understand patterns of human-caused wildfire ignition risk in the SW of Madrid, Central Spain, *Ecol. Modell.*, 221, 34–45, 2010.
- Ruiz-Mirazo, J., Martínez-Fernández, J., and Vega-García, C.: Pastoral wildfires in the Mediterranean: understanding their linkages to land cover patterns in managed landscapes, *J. Environ. Manage.*, 98, 43–50, 2012.
- Sá, A. C. L., Pereira, J. M. C., Charlton, M. E., Mota, B., Barbosa P. M., and Fotheringham, A. S.: The pyrogeography of sub-Saharan Africa: a study of the spatial non-stationarity of fire–environment relationships using GWR, *J. Geogr. Syst.*, 13, 227–248, doi:10.1007/s10109-010-0123-7, 2011.
- Sebastián-López, A., Salvador-Civil, R., Gonzalo-Jiménez, J., and San Miguel-Ayanz, J.: Integration of socio-economic and environmental variables for modelling long-term fire danger in Southern Europe, *European J. Forest Res.*, 127, 149–163, doi:10.1007/s10342-007-0191-5, 2008.



- Syphard, A. D., Radeloff, V. C., Keeley, J. E., Hawbaker, T. J., Clayton, M. K., Stewart, S. I., and Hammer, R. B.: Human influence on California fire regimes, *Ecol. Appl.*, 17, 1388–1402, 2007.
- Syphard, A. D., Radeloff, V. C., Keuler, N. S., Taylor, R. S., Hawbaker, T. J., Stewart, S. I., and Clayton, M. K.: Predicting spatial patterns of fire on a southern California landscape, *Int. J. Wildland Fire*, 17, 602–613, doi:10.1071/WF07087, 2008.
- SSTARS (Social Sciences Teaching and Research Statistics): Multicollinearity in Logistic Regression. Information Technology Computing Center, University of Kentucky, available at: <http://www.uky.edu/ComputingCenter/SSTARS/MulticollinearityinLogisticRegression.htm>, last access: 10 March 2012.
- Tulbure, M. G., Wimberly, M. C., Roy, D. P., and Henebry, G. M.: Spatial and temporal heterogeneity of agricultural fires in the central United States in relation to land cover and land use, *Landscape Ecol.*, 26, 211–224, doi:10.1007/s10980-010-9548-0, 2011.
- Vasconcelos, M. J. P., Silva, S., Tome, M., Alvim, M., and Pereira, J. M. C.: Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks, *Photogramm. Eng. Rem. S.*, 67, 73–8, 2001.
- Vazquez A. and Moreno J. M.: Patterns of lightning- and human-caused fires in peninsular Spain, *Int. J. Wildland Fire*, 8, 103–115, doi:10.1071/WF9980103, 1998.
- Vega-Garcia, C., Woodard, T., Adamowicz, W. L., and Lee, B.: A logit model for predicting the daily occurrence of human caused forest fires, *Int. J. Wildland Fire*, 5, 101–111, 1995.
- Vélez, R.: *La Defensa Contra Incendios Forestales. Fundamentos y Experiencias*. McGraw-Hill, Aravaca, Madrid, Spain, 2009.
- Viegas, D. X., Allgöwer, B., Koutsias, N., and Eftichidis, G.: Fire Spread and the Urban Wildland Interface Problem, edited by: Xanthopoulos, G., in: *Proceedings of the International Workshop on Forest Fires in the Wildland-Urban Interface and Rural Areas in Europe: an integral planning and management challenge*, MAICH, Chania, Greece, 22–34, 2003.
- Vilar, L., Woolford, D. G., Martell, D. L., and Martín, M. P.: A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain, *Int. J. Wildland Fire*, 19, 325–337, doi:10.1071/WF09030, 2010.
- Wang, Q., Ni, J., and Tenhunen, J.: Application of a Geographically-Weighted Regression Analysis to Estimate Net Primary Production of Chinese Forest, *Global Ecol. Biogeogr.*, 14, 379–393, 2005.
- Xu, D. Shao, G. Dai, L. Hao, Z. Tang, L., and Wang, H.: Mapping forest fire risk zones with spatial data and principal component analysis, *Sci. China Ser. E*, 49, 140–149, 2006.
- Yang, J., He, H. S., Shifley S. R., and Gustafson E. J.: Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands, *Forest Sci.*, 53, 1–15, 2007.