Predicting the occurrence of wildfires with binary structured additive regression models

Laura Ríos-Pena,^{a,*} Thomas Kneib,^b Carmen Cadarso-Suárez,^c Manuel Marey-Pérez^d

^aConservation Biology Department, Doñana Biological Station-CSIC, C/Américo Vespucio s/n, 41092 Seville, Spain.

^bChair of Statistics, Georg-August-University Göttingen, Humboldtalle 3, 37073 Göttingen, Germany.

^cUnit of Biostatistics, Department of Statistics and Operations Research, School of Medicine, University of Santiago de Compostela, C/San Francisco s/n, 15782 Santiago de Compostela, Spain.

^dPROEPLA Research Group, Polytechnic Institute, University of Santiago de Compostela, Campus Universitario s/n, 27002 Lugo, Spain.

*Corresponding author. E-mail: laura.rios@ebd.csic.es

Abstract

Wildfires are one of the main environmental problems facing societies today, and in the case of Galicia (north-west Spain), they are the main cause of forest destruction. This paper used binary structured additive regression (STAR) for modelling the occurrence of wildfires in Galicia. Binary STAR models are a recent contribution to the classical logistic regression and binary generalized additive models. Their main advantage lies in their flexibility for modelling non-linear effects, while simultaneously incorporating spatial and temporal variables directly, thereby making it possible to reveal possible relationships among the variables considered. The results showed that the occurrence of

wildfires depends on many covariates which display variable behaviour across space and time, and which largely determine the likelihood of ignition of a fire. The joint possibility of working on spatial scales with a resolution of 1×1 km cells and mapping predictions in a colour range makes STAR models a useful tool for plotting and predicting wildfire occurrence. Lastly, it will facilitate the development of fire behaviour models, which can be invaluable when it comes to drawing up fire-prevention and firefighting plans.

Key words: wildfires, structured additive regression models, covariates, voxel, penalized splines, Markov random fields.

1. Introduction

The concept of wildfire risk, defined as the probability of a wildfire occurring in an area within a given time interval, is well suited to strategic planning in fire and land management, since it integrates the probability of fire with its consequences (Bachmann and Allgower, 2000; Hardy 2005; Fairbrother and Turnley 2005; Finney et al. 2011; Scott 2006, Calkin et al. 2010). Quantitative approaches to risk assessment are still relatively new in wildland fire studies (Fairbrother and Turnley 2005; Finney et al. 2011; Scott 2006, Calkin et al. 2010). According to Miller and Ager (2013), two decision environments could be supported by fire risk analyses in the future, namely, those of fuel management and ignition management.

The 21st century has seen an increase in the number of wildfires (Plucinski 2013). Recent studies based on global burnt area estimate that 3.5 million km² are burnt every year (Chang and Song 2009; Giglio et al. 2010). The impact of climate change on

ignition occurrence has been documented by several authors (Kasischke and Turetsky 2006; Westerling et al. 2006), something that heightens the relevance of establishing future scenarios of fire risk conditions. Forest fires are the leading environmental problem in the Mediterranean region of Europe (Boubeta et al. 2015; Rodrigues et al. 2013; Fernandes et al. 2011; Fuentes-Santos et al. 2015; Rodríguez y Silva and González-Cabán 2010; Ruiz-Mirazo et al. 2012; Viedma et al. 2009; Vilar et al. 2010). In the study of wildfires, different concerns have been considered, such as the increase in forest area (Catry et al. 2009), overlap between urban development and forest (Lampin-Maillet et al. 2010; Mitsopoulos et al. 2014; Moreno 2014), new agrarian landscape patterns (Amatulli et al. 2007; Catry et al. 2009; Corcoran et al. 2007; Fernandes 2009; Kalabokidis et al. 2007; Martínez et al. 2009; Moreira et al. 2009, 2011; Verde and Zêzere 2010; Rego and Silva 2014), changes in traditional agrarian landscape patterns (Ruíz-Mirazo et al. 2012; Chas-Amil et al. 2013), agricultural abandonment (Acosta et al. 2005; Aranzabal et al. 2008; Falcucci et al. 2007), expansion of shrubland surface (Bajocco and Ricotta 2008; Díaz-Delgado et al. 2004; González et al. 2006; González and Pukkala 2007; Koutsias et al. 2012; Montané et al. 2009; Moreira et al. 2009; Mouillot et al. 2005; Nunes et al. 2005; Sebastian-Lopez et al. 2008), socio-economic changes (Arianoutsou 2001; Badia-Perpinya and Pallares-Barbera 2006; Cardille et al. 2001; Curt and Delcros 2010; Loboda and Csizar 2007; Martínez et al. 2009; Romero-Calcerrada et al. 2010; Vasconcelos et al. 2001; Vilar et al. 2010; Wittenberg and Malkinson 2009; De Zea Bermudez et al. 2009; Chas-Amil et al. 2013; Calviño et al. 2015), post-fire reactions (Conedera et al. 2011), and changes in income distribution and higher unemployment levels (González-Olabarria et al. 2015).

Different ways of modelling the risk of wildfires have been explored, depending upon

their origin (e.g., whether caused by lightning or arson) (Albert-Green et al. 2013). In recent years, a variety of statistical methods have been used to predict the occurrence of wildfires based on the presence/absence of fire, in addition to the logistic generalized linear models (logistic GLM, McGullagh and Nelder 1989) that have traditionally dominated the literature on ignition occurrence modelling (Serra et al. 2008; Woolford et al. 2011). In particular, models have been developed which explain the temporal variability in the likelihood of fires occurring over time. Rodrigues et al. (2014) applied the Mann-Kendall non-parametric test as well as the Spearman and Cox-Stuart tests for data at the *Nomenclature des Unités Territoriales Statistiques* (NUTS) 3 level. As it discriminates between clustering and stochastic sequences, Benavent-Corai et al. (2007) used the coefficient of variation (CV) to highlight the presence of clusters in the temporal process. The dataset that they studied consisted of forest fires recorded from 1 January 1989 to 31 December 2003 in the Valencian Region in Spain. Levin and Heimowitz (2012) examined statistical relationships (using Pearson's correlation coefficient) between weather variables and annual burnt areas. They also examined the effects of land-use and vegetation cover on fire frequency in Israel for all wildfires caused by humans in 2006 and 2010.

Many studies addressing optimal spatial management of forests under risk (Meilby et al. 2001) underscore the need to include spatial aspects in risk assessment (Koutsias and Karteris 1998; Koutsias 2003; Preisler et al. 2004). Kalabokidis et al. (2007) applied logistic regression and correspondence analysis to a spatial database which had been developed and managed in a GIS for the Sithonia peninsula in northern Greece. Artificial neural networks (ANNs, Zhanqing et al. 2001) have a great potential for modelling these events, since they can approximate mathematical relationships to which statistical models cannot be applied and in which relationships are not linear. Non-

normal distributions are common in wildfire data, and very high correlations are to be expected among landscape structure indices. To investigate these patterns, a study was made of the landscape structure and all pastoral wildfires recorded from 1988 to 2000 in 24 natural parks in Andalusia (Spain) (Ruiz-Mirazo et al. 2012). Viedma et al. (2009) used bootstrapping for validation of generalized linear models (GLMs) of the Sierra de Gredos range (Spain) for the period 1975-1990. Kens and Ager (2007) found quantitative risk assessment in the inclusion of stochastic spatial processes a major challenge. Tuia et al. (2008a) applied spatial clustering to fire data recorded from 1997 to 2003 in the Tuscany region of central Italy, using several independent indices (Voronoi polygon area index, Morishita index, and fractal dimension). In recent years, flexible regression models, such as generalized additive models (GAMs) (Hastie and Tibshirani 1990; Wood 2006), have been proposed for the purpose of studying possible non-linear effects of continuous covariates on the response variable (Roccaforte et al. 2012; Boer et al. 2009), i.e., these models support the fact that the response variable is modelled as a function of the smoothed predictor. Nevertheless, they have certain limitation when a spatial effect is included in the model.

Accordingly, this paper proposes the use of a structured additive regression (STAR) model with binary distribution, ultimately aimed at developing consistent, spatial-temporal models that can be easily integrated into a fire ignition-point prediction system.

2. Material and methods

2.1. Study Area

In terms of total wooded area, Spain is the third most forested state in Europe (with 15 million hectares), after Sweden and Finland (Forest Europe 2011). The Autonomous

Region (*comunidad autónoma*) of Galicia, where this study was conducted, is located in the north-west of the Iberian Peninsula. Galicia has a surface area of 29,574 km² (5.8% of mainland Spain) and a population of 2,795,422 (5.9% of the Spanish population), amounting to a population density of 99 persons/km², similar to the mean population density of Spain, i.e., 92 persons/km² (INE, 2014). To improve forest management, in 2008 Galicia was statutorily divided into 19 forest districts (Ley 43/2003, Decreto 43/2008).

The ratio of forested surface is notably higher than the national average (55.2%) (MAGRAMA 2014). As regards the local property regime, i.e., the basis on which holders are entitled to the land, most Galician forests are privately owned (97.2%) and are in the hands of a large number of individuals possessing an average of 2 ha per capita, which is in turn divided into an average of four plots (Marey-Pérez et al. 2012; Picos 2009). From a commercial point of view, Galicia is currently the leading timber forest product producer in Spain and the ninth in Europe (Balsa-Barreiro and Hermosilla, 2013). Hence, the timber industry is of significant economic importance. With annual revenues of €2,259,000,000 the forestry sector generates 26,000 direct and 50,000 indirect jobs and represents 4.5% of the regional GDP (FEARMAGA 2011).

Figure 1 shows the average annual distribution of the number of wildfires in Europe by province (Biot, 2009), and indicates that Galicia is one of the regions severely affected by wildfires (Comas et al. 2014). Indeed, between 1961 and 2014, over 251,106 wildfires were recorded in Galicia, affecting an estimated area of 1,830,000 ha, which accounts for approximately 66% of the region's surface area, being the accumulated burned area over the years higher than the current wooded forester's area of Galicia. Furthermore, it should be noted that many rural areas in the north-western part of Spain

have been abandoned and, as a consequence, agro-pastoral activities linked to the traditional rural lifestyle have vanished (Stellmes et al., 2013; Balsa-Barreiro and Hermosilla, 2013). Gómez-Vázquez et al. (2009) and Marey-Perez et al. (2010, 2014a, 2014b) analysed wildfires in relation to property rights, and found that collective private property rights, in particular, might be the cause of many fires. Caballero (2015) examined the process of change of ownership and land use management, role of government and causes of conflict as potential explanatory variables for wildfires. Given the importance of the problem, different authors have developed and validated methods designed to account for and predict fires in Galicia in recent years. In this study, we used binary STAR for modelling the occurrence of wildfires in Galicia.

2.2. Source and description of the datasets

For analysis purposes, we used the wildfire database of the Galician Regional Rural and Maritime Authority (*Consellería de Medio Rural y Mar*) (mediorural.xunta.gal). This database contains information on the site (geographical co-ordinates) and time (hour and date) at which a fire occurs. We selected wildfires recorded during the first half of August 2006 (see Figure 2). A total of 2,060 wildfires were recorded in that period: of these, 1,553 were due to arson, 222 to unknown causes, 198 to resurgence of a previous wildfire, 58 to other causes, 9 to power lines, 5 to smoking, 5 to bonfires, 3 to landfill leaks, 3 to agricultural burning, 3 to engines and machinery, and 1 to burning brush (Xunta de Galicia, 2014). The spatial component was made up of 30,685 cells, derived by dividing Galicia into a 1×1 km cells. For analysis and presentation of results, we used a map of Galicia divided into municipal areas, which are the closest to basic local administrative units and consist of a total of 315 municipalities (IGN 2011). The centroid of each cell is associated with a municipal area, so that each grid pertains to a municipality. This map enabled the effect of spatial correlation of municipalities to be

interpreted with respect to wildfire occurrence. Moreover, working on this scale made for easier interpretation of results and lower computational costs.

Daily weather covariates were obtained from meteorological monitoring stations distributed throughout the study area. A total of 71 meteorological monitoring stations were operational during the study period. The daily weather covariates selected in this study were: (i) average temperature; (ii) average relative humidity; and, (iii) rainfall. We further restricted our analysis to meteorological stations situated at a maximum distance of 30 km from any given wildfire. Due to differences in altitude between fires and the nearest meteorological stations, it was necessary to make a correction of the temperature and rainfall variables with respect to the difference in altitude. The variation in daily average temperature with height was $0.65^{\circ}\text{C}/100\text{m}$ to 11000 m (Nuñez and Colhoun 1986), i.e., $tH = 15^{\circ}\text{C} - 0.0065^{\circ}\text{C}/\text{m} \cdot H$ where tH is the air temperature at height , and H is the difference between the wildfire and the nearest meteorological station. Likewise, a simple method for correcting rainfall (P) in relation to differences in altimetry is as follows, $P_{(corrected)}(monthly) = P_{(monthly)} \pm 0.08$, i.e., the rainfall gradient amounted to an 8% variation per 100 m difference in height between any given ignition point and the meteorological station selected.

Rothermel fuel models (Rothermel, 1983) are a classification of different types of vegetation and their behaviour in fire. The combustibility of a forest system is defined as its capacity to burn, release enough material to be consumed and cause burning of neighbouring vegetation, thereby increasing the distribution and direction of wildfire. Each type of vegetation corresponds to a certain flammability and combustibility, which vary depending on the type and quantity of biomass and its spatial distribution or stratification (Galiana et al. 2009). The Rothermel fuel-model classification considered

13 types of fuel models, divided into 4 major groups based on the primary means of spreading fire, i.e., pasture, shrub, litter under trees and wood residuals. These models were drawn from the series of National Forest Inventories (MAGRAMA 2014). In addition, we included a fifth group, denoted as "unallocated", since Galician forest is extremely heterogeneous and this often makes it difficult to assign a forest area to one specific category. Table 1 provides a brief description of the response variable, explanatory covariates and spatial aggregation used for the purposes of our study.

2.3 Statistical methods

2.3.1. Structured Additive Regression

Structured additive regression (Fahrmeir et al. 2004) provides a generalization of generalized linear and additive models for a response from an exponential family, where a more general type of predictor including non-linear, spatial or random effects can be used to study the regression effects of a variety of covariates. In our analysis, we combined potentially non-linear effects of continuous covariates with a spatial effect based on the municipalities, as well as a random effect for such municipalities. The latter two effects capture any spatially structured and unstructured heterogeneity remaining after accounting for covariates.

In generic notation, the predictor of a STAR model for grid cell k at time t is given by

$$\eta_{kt} = \beta_0 + f_1(v_{kt}) + \dots + f_r(v_{kt}) \quad (1)$$

where β_0 is the overall intercept, and the functions $f_1, \dots, f_r$, represent regression effects of different types of covariates generically called v_{kt} . In our specific case, let p_{kt} denote the probability of occurrence of a wildfire at grid cell k at time t , with km^2 per day

being the unit measured. We then assume a logistic structured additive regression model

where

$$\log\left(\frac{p_{kt}}{1-p_{kt}}\right) = \eta_{kt} = \beta_0 + f_1(v_{kt}) + \dots + f_r(v_{kt}) + f_{struc}(v_{kt}) + f_{unstruc}(v_{kt})$$

In this model specification, the functions $f_1, \dots, f_r$ represent potentially non-linear effects of continuous covariates, while f_{str} and f_{unstr} represent spatially structured and unstructured heterogeneity for each municipality. The basic idea of STAR models is to approximate any of the covariate effects $f(v)$ by a linear combination of basis functions, such that

$$f(v) = \sum_{j=1}^J \beta_j B_j(v)$$

where $B_j(v)$ are the basis functions, while β_j denotes the corresponding basis coefficients. Specific choices relevant to our modelling exercise will be discussed in the next paragraph. In matrix notation, the STAR predictor (1) can now be rewritten as

$$\eta = \beta_0 \mathbf{1} + Z_1 \beta_1 + \dots + Z_r \beta_r$$

where η is obtained by stacking all individual observations, $\mathbf{1}$ refers to a vector of ones, Z_j arises from evaluating the basis functions at the observed covariates, and β_j is the vector of corresponding basis coefficients.

To ensure specific properties of the function estimates, such as smoothness in the case of non-linear, and spatial effects or shrinkage in the case of cell-specific random effects,

we relied on a Bayesian formulation where prior distributions are assigned to the vectors of regression coefficients. Generically, these are given by multivariate normal distributions.

$$p(\beta_j|\tau_j^2) \propto \exp\left(\frac{-1}{2\tau_j^2}\beta_j'K_j\beta_j\right)$$

where τ_j^2 is a prior variance determining the impact of the prior distribution on the function estimates, and K_j is the prior precision matrix that implements the smoothness/shrinkage assumptions. From a frequentist perspective, these prior distributions correspond to penalties of the form $\lambda_j\beta_j'K_j\beta_j$, where $\lambda_j \geq 0$ is the smoothing parameter that determines the impact of the penalty on the penalized likelihood estimate.

2.3.2. Model Components

For the non-linear effects of continuous covariates, we considered penalized splines (Eilers and Marx 1996, Brezger and Lang 2006), where a specific function $f(x)$ is approximated by a linear combination of B-spline basis functions. A common choice is cubic B-splines with 20 equidistant knots that yield a twice continuously differentiable (i.e., visually smooth) function estimate, while allowing for enough flexibility to capture a typical non-linear pattern. As a roughness penalty, squared first- or second-order differences are used, which correspond to the assumption of a first- or second-order random walk prior (with the latter being the default choice).

For the spatially structured effect, we used a Markov random field approach, where

each grid cell is associated with a separate regression effect (such that the basis functions correspond to indicator functions for the grid cells), while spatial smoothness is achieved by assuming an adjacency matrix structure for the penalty matrix K . In contrast, the same basis functions are used for the spatially unstructured part but in this case, the penalty matrix is reduced to an identity matrix which corresponds to the assumption of an independent identically distributed (i.i.d.) random effect. A more detailed discussion of potential model components and how they fit into the generic framework outlined above can be found in Fahrmeir et al. (2013).

2.3.3. Empirical Bayes Inference

In our analyses, we relied on an empirical Bayes approach that treats smoothness variances τ_j^2 as fixed unknowns to be estimated from the corresponding marginal posterior. Estimation is thus basically reduced to iterating between optimising a penalized logistic likelihood and numerical determination of the smoothing variances by approximate restricted maximum likelihood estimation (see Fahrmeir et. al. 2004, for details). The penalized likelihood estimates for the regression coefficients can be obtained by denoting the complete model as $\eta = Z\beta$, where $Z = [1, Z_1, \dots, Z_r]$ and $\beta' = (\beta_0, \beta'_1, \dots, \beta'_r)$, such that we iteratively compute

$$\hat{\beta}^{[t]} = (Z'W^{[t-1]}Z + K)^{-1}Z'W^{[t-1]}\tilde{y}^{[t-1]}$$

where $K = \text{blockdiag}(0, \lambda_1 K_1, \dots, \lambda_r K_r)$ and $W^{[t-1]}$ and $\tilde{y}^{[t-1]}$ are working weights and working observations as in standard generalized linear models. To derive the estimated model complexity of a specific term $f_j(v)$ in the structured additive predictor, one can consider the equivalent degrees of freedom df_j obtained from the trace of the

291 hat matrix

$$H = Z(Z'WZ + K)^{-1}Z'W$$

292

293 upon convergence of the penalized likelihood iterations. These degrees of freedom can
294 be interpreted as an effective parameter count, which can be used, for instance, to
295 quantify the deviation of a non-linear effect from the simple linear case, with one degree
296 of freedom for the slope parameter.

297

298 3. Results

299 We generated a model that studies the relationship between wildfires and daily weather
300 covariates, types of fuel model on the ground, and any possible spatial effects
301 corresponding to all the municipalities located in the study area. This model is
302 expressed as follows,

303

$$\begin{aligned} \log\left(\frac{p_{kt}}{1 - p_{kt}}\right) = & \beta_0 + f_1(\text{altitude}_k) + f_2(\text{ta}_{\text{mean}_k}) + f_3(\text{hr}_k) + f_4(\text{pp}_k) \\ & + f_5(\text{unallocated}_k) + f_6(\text{pastures}_k) + f_7(\text{shrub}_k) + f_8(\text{litter}_k) \\ & + f_9(\text{wood}_{\text{residuals}_k}) + f_{\text{struc}}(\text{cdconc}_k) + f_{\text{unstruc}}(\text{cdconc}_k) \end{aligned}$$

304

305 where k is the index of the set of grid cells in the study area and period, P_{kt} denotes the
306 probability of ignition point in the voxel k at time t , given the observed covariate set.
307 The functions f_j are smooth effects capturing the potentially non-linear relationship
308 between the explanatory variables and the logit probability of the response variable.
309 These functions were estimated using Bayesian cubic P-splines with a second-order
310 random walk prior and 20 inner knots. The spatial effects were estimated using Markov
311 Random Fields (MRF, f_{str}) and using i.i.d. Gaussian random effects (f_{unstr}).

To demonstrate that the structured additive regression with binary response improves predictions of wildfires, as compared to classical logistic regression models and generalized additive models, a table was drawn up showing the results yielded by the following models: GLM; GLM with spatial effect; GAM; and lastly, STAR. The distribution of the response variable was binary in all cases. The selection criteria used to choose the model that produced the best estimate were the Akaike Information Criterion (AIC) and Generalized Cross-Validation (GCV), which are shown in Table 2 below, along with the $-2 \times \log$ -likelihood and degrees of freedom (df).

As can be seen from Table 2, the model that best estimated the occurrence of wildfires was the STAR model with binary response. Accordingly, a breakdown of the results obtained from the latter is now given in Figure 3 below.

The results of the smoothed effects of the weather and fuel model covariates are shown in Figure 3. The effect of altitude (Figure 3a) indicates a decrease in wildfire probability as altitude increases. This is reasonable, since most of the wildfires occurred in coastal areas, where altitudes are lower. The estimated degrees of freedom for this covariate were 2,502 which indicates slight non-linearity, with a gradual downward trend at low altitudes and a sharp decrease as altitude increases. Across the study period, the occurrence of fires in the region would thus appear to have been related to lower altitudes. It is noteworthy, however, that the credible bands increased in width from an altitude of approximately 1000 m, meaning that these results should not be overinterpreted, since there were few data showing altitudes above 1000 m in the study area.

337

338 Average daily temperature (Figure 3b) appeared to have a non-linear effect, implying
339 strongly increasing occurrence probabilities up to a level of around 20°C. From this
340 temperature upwards, the probability of wildfire occurrence remained approximately
341 constant up to 30°C. From 30°C upwards, the probability started to increase again,
342 though not as sharply as at low temperatures (<20°C). The degrees of freedom took a
343 value of 4,069, which is in accordance with the visual impression of a non-linear effect.
344 The effect of relative humidity (Figure 3c) was rising for low relative humidity values
345 up to about 35%: for higher values, the overall effect was downward, with the sharpest
346 decline being observed for high humidity of around 75%-80%. Wildfire occurrence was
347 thus associated with low relative humidity, which is reasonable in the light of the fact
348 that most of the wildfires occurred in coastal areas, where relative humidity below 50%
349 is common. Precipitation (Figure 3d) had an almost linearly decreasing effect.

350

351 In terms of type of fuel model, the effect of the unallocated category (Figure 3e) was
352 non-linear and increasing, i.e., the greater the surface area with no known type of plant
353 cover, the greater the likelihood of wildfire. The effect of pasture (Figure 3f) was linear
354 and increasing. The effect of shrub (Figure 3g) showed that increasing the amount of
355 area in this category also increased the probability of wildfires occurring; this category
356 had a strong presence throughout the study area (see Appendix). The effect of litter
357 (Figure 3h) was non-linear, with an increasing effect up to values of around 400,000 m²,
358 after which the effect remained stable up to 1 million m². The effect of wood residuals
359 (Figure 3h) was linear and almost constant. However, since there was a low overall
360 concentration of wood residuals/m², the results for this covariate should not be
361 overinterpreted.

The effect of the spatial covariate (cdconc, corresponding to municipalities) was divided into spatially structured (Figure 4a and 4b) and unstructured effects (Figure 4c and 4d), as mentioned above. The plot with structured spatial effects (Figure 4a) shows that municipalities situated in the north and north-east had a low occurrence of fires as compared to municipalities situated in the west and south, after taking covariates into account. While 95% credible intervals (Figure 4b) present in black municipalities denote that the effect was significantly negative, white corresponds to significantly positive effects, and grey indicates a lack of significance.

The results of the unstructured spatial effects (Figures 4c and 4d) were mostly non-significant. Municipalities shown in pink had a greater possibility of a local effect than did those drawn in green. The Chandrexa de Queixa municipal area (depicted in the south-east in black) calls for a more detailed study to ascertain the possible causes of the singular pattern of wildfire in this area.

4. Discussion and conclusions

For over 40 years, researchers and forest services offices have been developing models that would be capable of predicting fire occurrence zones (Koutsias et al. 2015; Martínez-Fernández et al. 2013; Koutsias et al. 2010). This process has faced two major difficulties relating, firstly, to the randomness of the phenomenon, especially in the case of arson, and secondly, to the availability of statistics and computer tools that would enable accurate models to be developed for a given spatial and temporal context.

The designated goal of our study was to evaluate the predictability of ignition points by

using flexible regression techniques based on STAR models, a methodology that has been shown to provide greater accuracy than the simpler special cases of logistic GLMs and GAMs. Additive or geo-additive models with non-linear interaction effects, which consider spatial information in the form of a location variable, or models with random effects, can accelerate the development of fire behaviour models and prove most useful for drawing up fire prevention and firefighting plans. Their principal advantage lies in their flexibility, in that they can include spatial and temporal covariates along with the potentially non-linear effects of other continuous covariates. Moreover, these models generate maps of both structured and unstructured effects that can be plotted separately (Fahrmeir et al. 2013).

The validation process performed with ignition points across Galicia, a region with an arson problem of major dimensions, working at spatial scales with a pixel resolution of 1×1 km, a temporal resolution of one day, and physiographic, climatic and land-use covariates, combined with the possibility of mapping the predictions in a colour range, shows binary STAR models to be an important tool for predicting and managing wildfires in this region.

With respect to the model variables and results, several authors have studied wildfires from an environmental perspective for many years (Vega et al. 2009b; Lombao et al. 2014), and there are also various studies on the causes and consequences of emergency situations (González-Alonso and Merino-de-Miguel 2009; Balsa-Barreiro and Hermosilla 2013). In addition, while many authors have considered physiographic variables in their analyses, only a few have analysed the influence of altitude on fire ignition points. For instance, Vilar et al. (2010) reported a decrease in fire risk

probability with increasing altitude in the Madrid Region. This effect could be explained by less human land-use and activity at high elevations. Ordóñez et al. (2012) also found that the increase in altitude minimized wildfire risk in the Region of Leon. Our results appear to agree with these results, showing that increased human presence in near-shore areas also increases the risk of arson-induced wildfires.

Weather variables have been considered by most studies addressing all fire hazards. While certain temperature and humidity conditions are essential for wildfire to spread, in the case of arson-induced fires, these logically fail to explain the cause. Studies conducted in Galicia, such as that by Molina-Terrén and Cardil (2015), have studied the influence of temperature on large fires. Chas-Amil et al. (2015) considered precipitation and temperature in an attempt to understand the reasons underlying wildfire occurrences in 3600 parishes (the *parróquia* or administrative parish being the lowest tier of territorial designation in Galicia). Their results show that, whereas average temperature and precipitation are important when it comes to explaining the presence of fire, mean precipitation is not significant in the case of arson-induced fires. Likewise focusing on Galicia, Barreal and Loureiro (2015) used average summer rainfall and maximum summer temperature in Moran's and LISA statistics to analyse the spatial patterns of wildfires. Boubeta et al. (2015) introduced days without rain as a meteorological variable in a Poisson regression model for the entire region, and it was this model which best explained the concentration of wildfires in some areas. Our results appear to be in line with those of other authors, inasmuch as the presence of fire depends on certain weather conditions that often occur in Galicia in summer and facilitate incendiary activity.

Land-use variables, analysed statically or dynamically, have been commonly used to explain fire risk. Two main types of analysis have traditionally been used: one seeks to ascertain how changes in land-use affect the risk of fire; the other seeks to ascertain which type of land-use can explain an incendiary crisis. We adopted the latter approach. Much like us, when studying the occurrence of large fires in the south of France, Gantaume and Jappiot (2013) found that shrubland and pasture tended to appear in areas with fire risk. As explained by Vélez (2000) and Marino et al. (2014), shrubland and pasture are often associated with the practice of burning to rejuvenate pasture land in Spain. This means that the presence of shrub could well be a consequence of the form of land management rather than being genuinely associated with the occurrence of wildfires. Bisquert et al. (2014) constructed a fire danger index for Galicia and the neighbouring Region of Asturias. In the results, they note that increases in shrub and grassland areas were found to be significant in the accuracy of the index. Information relating to litter under trees and wood residuals are integrated into fuel-model types. Aguado et al. (2007) studied the characteristics and moisture content of forest residuals and considered them critical to fire risk. Chuvieco et al. (2010) analysed which variables influenced risk of fire in three regions of Spain (Aragon, Madrid and Valencia), and found dead-fuel moisture content to be an important risk factor. In their fire risk index for the Valencian Region, Vicente and Crespo (2012) introduced fuel conditions in the group of variables denoted as "ignition danger" and increased the accuracy of the index.

In terms of spatial distribution of ignition points, Comas et al. (2014) identified the south-western corner of Galicia as being especially important, a finding that is in good agreement with the area shown by our model to contain spatial aggregation points and

spatio-temporal structures with distances in a range of 0-12 km. In the north-west of Galicia, Fuentes-Santos et al. (2013) established spatial dependence between ignition points and an interaction radius of 3-4 km, corresponding to a lower concentration than in our results. Considering the region as a whole, Chas-Amil et al. (2015) noted that the highest concentrations were produced both in areas of higher population density, a finding that coincides with our results, and areas of greater depopulation. The difference in results might, at least in part, be related to the different observation times: August 2006 was an exceptional case in the history of fires in Galicia.

In summary, we believe logistic STAR models to be an important contribution to the toolbox of forestry managers and statisticians. Future challenges to the improvement of STAR models and their application to wildfire prediction are adding new variables at the scale of pixels used here. It should be noted that, in some cases, the estimates of the effect of certain covariates, such as rainfall, pasture group, litter-under-trees group and wood-residuals group, were seen to be perfectly linear. Improved data computation, the incorporation of variables linked to forest fires and a new statistical methodology would thus constitute a solid line of future research.

Appendix

1. Dataset

The total of nine variables studied can be classified into two types:

- Meteorological covariates. Data were collected via meteorological stations (see Figure 5). During the first half of August 2006, a total of 71 meteorological stations were in operation. Data were only taken into account where the distance between the

meteorological station and the wildfire was less than 30 km. The area from which data were extracted is shown in Figure 4.

- Rothermel fuel-model groups. Figure 6 shows the spatial distribution of the five fuel-model types in Galicia.

2. Software

All statistical analyses were performed using the R statistical software environment, and the following packages in particular:

- BayesX. This package provides functionality for the purpose of exploring and visualising the results of structured additive regression models obtained with BayesX. It also provides functions that read, write, and manipulate objects required to visualize spatial effects obtained with BayesX.

- R2BayesX. This package provides an interface between R and the BayesX statistical software (<http://www.BayesX.org>). To analyse the occurrence of wildfires, we used the `bayesx` function, which is the main function for fitting models in the R2BayesX package.

Acknowledgements

T. Kneib and C. Cadarso-Suárez would like to thank the Spanish Ministry of Innovation & Competitiveness for support from the MTM2014-52975-C2-1-R project; and M. Marey-Pérez and C. Cadarso-Suárez would like thank the Galician Regional Authority (*Xunta de Galicia*) for supporting research of the INBIOEST network (2014-PG141) and Competitive Research Reference Group (PROePLA CN2012/015).

References

- Albert-Green, A., Dean, C.B., Martell, D.L., Woolford, D.G., 2013. A Methodology for Investigating Trends in Changes in the Timing of the Fire Season with Applications to Lightning-Caused Forest Fires in Alberta and Ontario, Canada. *Canadian Journal of Forest Research* 43, 39–45.
- Acosta, A., Carranza, M.L., Giancola, M., 2005. Landscape change and ecosystem classification in a municipal district of a small city (Isernia, Central Italy). *Environmental Monitoring and Assessment* 108, 323-335.
- Aguado, I., Chuvieco, E., Boren, R., Nieto, H., 2007. Estimation of dead fuel moisture content from meteorological data in Mediterranean areas. Applications in fire danger assessment. *International Journal of Wildland Fire* 16, 390-397.
- Amatulli, G., Pérez-Cabello, F., de la Riva, J., 2007. Mapping lightning/human-caused wildfire occurrence under ignition point location uncertainty. *Ecological Modelling* 200:321-333.
- Aranzabal, I., Schmitz, M.F., Aguilera, P., Pineda, F.D., 2008. Modelling of landscape changes derived from the dynamics of socio-ecological systems: a case of study in a semiarid Mediterranean landscape. *Ecological Indicators* 8, 672-685.
- Arianoutsou, M., 2001. Landscape changes in Mediterranean ecosystems of Greece: Implications for fire and biodiversity issues. *Journal of Mediterranean Ecology* 2, 165-

178.

Bachmann, A., Allgower, B., 2000. The need for a consistent wildfire risk terminology. In: Neuenschwander, L., Ryan, K., Gollberg, G. (Eds.), Crossing the Millennium: Integrating Spatial Technologies and Ecological Principles for a New Age in Fire Management. The University of Idaho and the International Association of Wildland Fire, Moscow, ID, pp. 67-77.

Badia-Perpinya, A., Pallares-Barbera, M., 2006. Spatial distribution of ignitions in Mediterranean periurban and rural areas: the case of Catalonia. International Journal of Wildland Fire 15, 187-196.

Bajocco, S., Ricotta, C., 2008. Evidence of selective burning in Sardinia (Italy): Which land-cover classes do wildfires prefer? Landscape Ecology 23, 241-248.

Balsa-Barreiro, J., Hermosilla, T., 2013. Socio-geographic analysis of wildland fires: causes of the 2006's wildfires in Galicia (Spain). Forest Systems 22(3): 497-509.

Barreal, J., Loureiro, M.L., 2015. Modelling Spatial Patterns and Temporal Trends of Wildfires in Galicia (NW Spain). Forest Systems, Volume 24, Issue 2, e-022, xx pages.

<http://dx.doi.org/10.5424/fs/2015242-05713>.

Benavent-Corai, J., Rojo, C., Suárez-Torres, J., Velasco-García, L., 2007. Scaling properties in forest fire sequences: the human role in the order of nature. Ecol Model 205: 336–342.

562 Bisquert, M., Sánchez, J.M., Caselles, V., 2014. Modelling Fire Danger in Galicia and
 563 Asturias (Spain) from MODIS Images. *Remote Sens* 6, 540-554.
 564
 565 Birot, Y., 2009. Living with wildfires: What science can tell us. *EFI Discussion* 15,
 566 European Forest Institute.
 567
 568 Boer, M.M., Sadler, R.J., Wittkuhn, R.S., McCaw, L., Grierson, P.F., 2009. Long-term
 569 impacts of prescribed burning on regional extent and incidence of wildfires-Evidence
 570 from 50 years of active fire management in SW Australian forests. *Forest Ecology and*
 571 *Management* 259: 132–142.
 572
 573 Boubeta, M., Lombardía, M.J., Marey-Pérez, M.F., Morales, D., 2015. Prediction of
 574 Forest Fires Occurrences with Area-Level Poisson Mixed Models. *Journal of*
 575 *Environmental Management* 154: 151–58.
 576
 577 Brezger, A., Lang, S., 2006. Generalized additive regression based on Bayesian
 578 Psplines. *Computational Statistics and Data Analysis* 50, 967–991.
 579
 580 Calkin, D.E., Ager, A.A., Gilbertson-Day, J., Scott, J.H., Finney, M.A., Schrader-
 581 Patton, C., Quigley, T.M., Strittholt, J.R., Kaiden, J.D., 2010. Wildfire risk and hazard:
 582 procedures for the first approximation. U.S. Forest Service, Rocky Mountain Research
 583 Station.
 584
 585 Calviño, M., Chas-Amil, M., García-Martínez, E., Touza, J., 2015. Wildfire risk
 586 associate with land cover types, topography and human activities within and outside

587 wildland-urban interface. *Landscape Ecology* (submitted).

588

589 Cardille, J.A., Ventura, S.J., Tuner, M.G., 2001. Environmental and social factors
590 influencing wildfires in the Upper Midwest, USA. *Ecological Applications* 11, 111-127.

591 Catry, F.X., Rego, F.C., Bação, F.F., Moreira, F., 2009. Modeling and Mapping
592 Wildfire Ignition Risk in Portugal. *International Journal of Wildland Fire* 18: 921–31.

593

594 Chang, D., Song, Y., 2009. Comparison of L3JRC and MODIS global burned area
595 products from 2000 to 2007, *J. Geophys. Res.*, 114, D16106,
596 doi:10.1029/2008JD011361.

597

598 Chas-Amil, M.L., Prestemon, J., McClean, C.J., Touza, J., 2015. Human-ignited
599 wildfire patterns and responses to policy shifts. *Applied Geography* 56, 164-176.

600

601 Chas-Amil, M.L., Touza, J., García-Martínez, E., 2013. Forest fires in the wildland-
602 urban interface: a spatial analysis of forest fragmentation and human impacts. *Applied*
603 *Geography* 43, 127-137.

604

605 Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M.P., Vilar, L.,
606 Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J.R.,
607 Herrera, M.A., Zamora, R., 2010. Development of a framework for fire assessment
608 using remote sensing and geographic information system technologies. *Ecological*
609 *Modelling* 221, 46-58.

610

611 Comas, C., Costafreda-Aumedes, S., Vega-García, C., 2014. Characterizing

612 configurations of fire ignition points through spatiotemporal point processes. *Natural*
 613 *Hazards and Earth System Sciences* 2, 2891-2911.
 614
 615 Conedera, M., Torriani, D., Neff, C., Ricotta, C., Bajocco, S., Pezzatti, G.B., 2011.
 616 Using Monte Carlo simulations to estimate relative fire ignition danger in a low-to-
 617 medium fire-prone region. *For. Ecol. Manage.* 261: 2179-2187.
 618
 619 Corcoran, J.G., Higgs, C.B., Ware, P.N., 2007. The Use of Spatial Analytical
 620 Techniques to Explore Patterns of Fire Incidence: A South Wales Case Study.
 621 *Computers, Environment and Urban Systems* 31, 623-647.
 622
 623 Curt, T., Delcros, P., 2010. Managing road corridors to limit fire hazard. A simulation
 624 approach in southern France. *Ecological Engineering* 4, 1-12.
 625
 626 De Zea Bermudez, P., Mendes, J., Pereira, J.M.C., Turkman, K.F., Vasconcelos, M.J.P.,
 627 2009. Spatial and temporal extremes of wildfire sizes in Portugal (1984–2004), *Int. J.*
 628 *Wildland Fire*, 18, 983–991.
 629
 630 Díaz-Delgado, R., Lloret, F., Pons, X., 2004. Spatial patterns of fire occurrence in
 631 Catalonia, NE, Spain. *Landscape Ecology*, 731-745.
 632
 633 Eilers, P.H., Marx, B.D., 1996. Flexible smoothing with B-splines and Penalties.
 634 *Statistical Science* 11, 89-121.
 635
 636 Fahrmeir, L., Kneib, T., Lang, S., 2004. Penalized additive regression for spacetime

637 data: a Bayesian perspective. *Statistica Sinica* 14, 731-761.

638

639 Fahrmeir, L., Kneib, T., Lang, S., Marx, B., 2013. *Regression - Models, Methods and*

640 *Applications*, Springer.

641

642 Fairbrother, A., Turnley, J.G., 2005. Predicting Risks of Uncharacteristic Wildfires:

643 Application of the Risk Assessment Process. *Forest Ecology and Management* 211(1-2):

644 28–35.

645

646 Falcucci, A., Maiorano, L., Boitani, L., 2007. Changes in land-use/land-cover patterns

647 in Italy and their implications for biodiversity conservation. *Landscape Ecology* 22,

648 617-631.

649

650 Federación de Aserraderos y Rematantes de Madera de Galicia (FEARMAGA), 2011.

651 Resultados Industria de la Madera de Galicia 2010. Santiago de Compostela.

652

653 Fernandes, P.M., Rego, F.C., Rigolot, E., 2011. The FIRE PARADOX Project: Towards

654 Science-Based Fire Management in Europe. *Forest Ecology and Management* 261(12):

655 2177–78.

656

657 Fernandes, P.M., Botelho, H.S., Rego, F.C., Loureiro, C., 2009. Empirical modelling of

658 surface fire behaviour in maritime pine stands. *International Journal of Wildland Fire*

659 18, 698-710.

660

661 Finney, M., McHugh, C.W., Grenfell, I.C., Riley, K.L., Short, K.C., 2011. A Simulation

of Probabilistic Wildfire Risk Components for the Continental United States. *Stochastic Environmental Research and Risk Assessment* 25(7): 973–1000.

Forest Europe, UNECE, FAO, 2011. State of Europe's Forests 2011. Status and Trends in Sustainable Forest Management in Europe. Instituto Nacional de Estadística (INE). National Statistics Institute, 2013-2014. Available at <http://www.ine.es> [30 May 2014].

Fuentes-Santos, I., González-Manteiga, W., Mateu, J., 2015. Consistent Smooth Bootstrap Kernel Intensity Estimation for Inhomogeneous Spatial Poisson Point Processes. *Scand J Statist*, doi: 10.1111/sjos.12183.

Galiana Martín, L., Aragonese Domínguez, C., Montiel Molina, C., De la Cita Benito, F., Fernández Ramiro, M., 2009. Caracterización de los escenarios del fuego en España, en 5º Congreso Forestal Nacional. Montes y sociedad: saber qué hacer. S.E.C.F.-Junta de Castilla y León, Ávila, 21-15 de septiembre de 2009. Ref. 5CFE01-450.

Gantaume, A., Jappiot, M., 2013. What causes large fires in Southern France. *Forest Ecology and Management* 294, 76-85.

Giglio, L., Randerson, J.T., van der Werf, G.R., Kasibhatla, P.S., Collatz, G.J., Morton, D.C., DeFries, R.S., 2010. Assessing Variability and Long-Term Trends in Burned Area by Merging Multiple Satellite Fire Products. *Biogeosciences* 7(2008): 1171-86.

Gómez-Vázquez, I., Álvarez-Álvarez, P., Marey-Pérez, M., 2009. Conflicts as enhancers or barriers to the management of privately owned common land: a method to

analyze the role of conflicts on a regional basis. *For. Policy Econ.*, 11 (2009), pp. 617-627.

González-Alonso, F., Merino-de-Miguel, S., 2009. Integration of AWiFS and MODIS active fire data for burn mapping at regional level using the Burned Area Synergic Algorithm (BASA). *International Journal of Wildland Fire* 18, 404-414.

González, J.R., Palahi, M., Trasobares, A., Pukkala, T., 2006. A fire probability model for forest stands in Catalonia (north-east Spain). *Annals of Forest Science* 63, 169-176.

González, J.R., Pukkala, T., 2007. Characterization of forest fire in Catalonia (northeast Spain). *European Journal of Forest Research* 126, 421-429.

González-Olabarria, J.R., Mola-Yudego, B., Coll, L., 2015. Different factors for different causes: analysis of the spatial aggregation of fire ignitions in Catalonia (Spain). *Risk analysis*. 35 (7):1197-1209.

Hardy, C.C., 2005. Wildland fire hazard and risk: Problems, definitions, and context. *Forest Ecology and Management* 211, 73-82.

Hernandez-Magallanes I (2014) Combining Ungrouped and Grouped Wildfire Data to Estimate Fire Risk. *Environmetrics* 25(6): 365–83.

Instituto Nacional de Estadística (INE), 2014. Censo de Población de España 2011. Madrid.

712 Kalabokidis, K.D., Koutsias, N., Konstantinidis, P., Vasilakos, C., 2007. Multivariate
 713 analysis of landscape wildfire dynamics in a Mediterranean ecosystem of Greece. *Area*
 714 39, 392-402.
 715
 716 Kasischke, E.S., Turetsky, M.R., 2006. Recent Changes in the Fire Regime across the
 717 North American Boreal region - Spatial and Temporal Patterns of Burning across
 718 Canada and Alaska. *Geophysical Research Letters* 33(9): L09703.
 719
 720 Kerns, B.K., Ager, A., 2007. Risk assessment for biodiversity conservation planning in
 721 Pacific Northwest forests. *Forest Ecology and Management* 246, 38–44.
 722
 723 Koutsias, N., Arianoutsou, M., Kallimanis, A. S., Mallinis, G., Halley, J. M., &
 724 Dimopoulos, P., 2012. Where did the fires burn in Peloponnisos, Greece the summer of
 725 2007? Evidence for a synergy of fuel and weather. *Agricultural and Forest Meteorology*,
 726 156, 41-53.
 727
 728 Koutsias, N., Allgöwer, B., Kalabokidis, K., Mallinis, G., Balatsos, P., Goldammer,
 729 J.G., 2015. Fire occurrence zoning from local to global scale in the European
 730 Mediterranean basin: implications for multi-scale fire management and policy. *iForest*
 731 10.3832/ifor1513-008.
 732
 733 Lampin-Maillet, C., Mantzavelas, A., Galiana, L., Jappiot, M., Long, M., Herrero, G.,
 734 Karlsson, O., Iossifina, A.T., Thanassis, P., 2010. Wildland urban interfaces, fire
 735 behaviour and vulnerability: characterization, mapping and assessment. In: *Towards*
 736 *Integrated Fire Management. Outcomes of the European Project Fire Paradox Finland:*

737 European Forest Institute. Pp 71-92.

738

739 Levin, N., Heimowitz, A., 2012. Mapping spatial and temporal patterns of

740 Mediterranean wildfires from MODIS. *Remote Sensing of Environment* 126, 12-26.

741

742 Li, Z., Khananian, A., Fraser, R., Cihlar, J., 2001. Automatic detection of fire smoke

743 using artificial neural networks and threshold approaches applied to AVHRR imagery,

744 *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 9.

745

746 Loboda, T., Csiszar, I., 2007. Reconstruction of Fire Spread within Wildland Fire

747 Events in Northern Eurasia from the MODIS Active Fire Product. *Global and Planetary*

748 *Change*, 56(3-4): 258-273.

749

750 Lombao, A., Barreiro, A., Carballas, T., Fontúrbel, M.T., Martín, A., Vega, J.A.,

751 Fernández, C., Díaz-Raviña, M., 2014. Changes in soil properties after a wildfire in

752 Fragas do Eume Natural Park (Galicia, NW Spain). *Catena* (in press).

753 doi:10.1016/j.catena.2014.08.007.

754

755 Marey-Pérez, M.F., Gómez-Vázquez, I., 2010. Modelo para la caracterización del nivel

756 de participación social y la conflictividad en los Montes Vecinales en Mano Común

757 (MVMC) en Galicia. *Span. J. Rural Dev.*, 1 (2010), pp. 85–102.

758

759 Marey-Perez, M.F., Rodriguez-Vicente, V., Alvarez-Lopez, J.C., 2012. Practical

760 Application of Multivariant Analysis Techniques to the Forest Management of Active

761 Farmers in the Northwest of Spain. *Small-Scale Forestry* 11, 453-476.

762

763 Marino, E., Hernando, C., Planelles, R., Madrigal, J., Guijarro, M., Sebastián, A., 2014.

764 Forest fuel management for wildfire prevention in Spain: a quantitative SWOT analysis.

765 International Journal of Wildland Fire 23, 373-384.

766

767 Martínez, J., Vega-García, C., Chivieco, E., 2009. Human-caused wildfire risk rating for

768 prevention planning in Spain. Journal Environmental Management 90, 1241–1252.

769

770 McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. Chapman and Hall.

771 Meilby, H., Strange, N., Thorsen, B.J., 2001. Optimal spatial harvest planning under

772 risk of windthrow. Forest Ecology and Management 149, 15-31.

773

774 Miller, C., Ager, A.A., 2013. A Review of Recent Advances in Risk Analysis for

775 Wildfire Management. International Journal of Wildland Fire 22(1): 1–14.

776

777 Ministry of Agriculture, 2014. IV National Forest Inventory. Available at:

778 [http://www.magrama.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-](http://www.magrama.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/)

779 [forestal-nacional/](http://www.magrama.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/) [30 May 2014].

780

781 Ministry of Agriculture, 2014. IV National Forest Inventory. Available at:

782 http://www.magrama.gob.es/es/biodiversidad/estadisticas/Incendios_default.aspx#para3

783 [30 May 2014].

784

785 Ministry of Agriculture, Food and Environment of the Spain Government

786 (MAGRAMA), 2014. IV National Forest Inventory.

787

788 Mitsopoulos, I., Mallinis, G., Karali, A., Giannakopoulos, C., Arianoutsou, M., 2014.
789 Mapping fire behaviour in a Mediterranean landscape under different future climate
790 change scenarios. In: International Conference ADAPT to CLIMATE, Nicosia, Cyprus.
791 27-28 March 2014.

792

793 Montané, F., Casals, P., Taull, M., Lambert, B., Dale, M., 2009. Spatial patterns of
794 shrub cover after different fire disturbances in the Pyrenees. *Ann. For. Sci.*, 66, 612–
795 620.

796

797 Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, E., Barbati,
798 A., Corona, P., Vaz, P., Xanthopoulos, G., Mouillot, F., Bilgili, E., 2011. Landscape
799 wildfire interactions in southern Europe: implications for landscape management.
800 *Journal of Environmental Management* 92, 2389-2402.

801

802 Moreira, F., Vaz, P., Catry, F., Silva, J.S., 2009. Regional variations in wildfire
803 susceptibility of land-cover types in Portugal: implications for landscape management
804 to minimize fire hazard. *International Journal of Wildland Fire* 18, 563-574.

805

806 Moreno, J.M., (Ed.) 2014. FUME. Forest fires under climate, social and economic
807 changes in Europe, the Mediterranean and other fire-affected areas of the world.
808 Lessons learnt and Outlook. Col. Book. 32p.

809

810 Mouillot, F., Ratte, J., Joffre, R., Mouillot, D., Rambal, S., 2005. Long-term forest
811 dynamic after land abandonment in a fire prone Mediterranean landscape (central

812 Corsica, France). *Landscape Ecology* 20, 101-112.

813

814 Nunes, M.C.S., Vasconcelos, M.J., Pereira, J.M.C., Dasgupta, N., Alldredge, R.J., 2005.

815 Land cover type and fire in Portugal: Do fires burn land cover selectively? *Landscape*

816 *Ecology* 20, 661-673.

817

818 Núñez, E.A., Colhoun, A., 1986. Note on air temperature lapse rates on Mount

819 Wellington. *Pap. Proc. R. Soc. Tasmania* 120, 11-15.

820

821 Ordóñez, C., Saavedra, A., Rodríguez-Pérez, J.R., Castedo-Dorado, F., Covián, E.,

822 2012. Using model-based geostatistics to predict lightning-caused wildfires.

823 *Environmental Modelling & Software* 29, 44- 50.

824

825 Pereira, P., Cerdá, A., Úbeda, X., Mataix-Solera, J., Arcenegui, V., Zavala, L.M., 2015.

826 Modelling the Impacts of Wildfire on Ash Thickness in a Short-Term Period. *Land*

827 *Degrad. Develop.*, 26, 180–192.

828

829 Picos, J., 2009. El sector forestal en Galicia: problemáticas y alternativas de futuro.

830 Jornadas sobre políticas públicas en tiempos de crisis: austeridad, eficacia y fomento de

831 la economía de Galicia. Galician School of Public Administration, Santiago de

832 Compostela, Spain.

833

834 Plucinski, M.P., 2013. Modelling the Probability of Australian Grassfires Escaping

835 Initial Attack to Aid Deployment Decisions. *International Journal of Wildland Fire*

836 22(4): 459–68.

837

838 Preisler, H.G., Brillinger, D.R., Burgan, R.E., Beniot, J.W., 2004. Probability based
839 models for estimation of wildfire risk. *International Journal of Wildland Fire* 13, 133-
840 142.

841

842 Rego, F.C., Silva, J.S., 2014. Wildfires and landscape dynamics in Portugal: a regional
843 assessment global implications in: J.C. Azevedo et al. (eds.), *Forest Landscapes and*
844 *Global Change: 51 Challenges for Research and Management*. New York.

845

846 Roccaforte, J.P., Fule, P.Z., Chancellor, W.W., Laughlin, D.C., 2012. Woody debris and
847 tree regeneration dynamics following severe wildfires in Arizona ponderosa pine
848 forests. *Can. J. For. Res.-Revue Can. De Rec. For.*, 42, 593–604.

849

850 Rodrigues, M., de la Riva, J., Fotheringham, S., 2014. Modeling the spatial variation of
851 the explanatory factors of human-caused wildfires in Spain using geographically
852 weighted logistic regression. *Applied Geography* 48, 52-63.

853

854 Rodríguez, P., Anca, A., Diéguez, A., Arenas, S., Amann, G., 2013. A necesidade
855 dunha xestión eficiente do monte galego. *Grial Dialogos*, 3, 26–35.

856

857 Rodríguez-Silva, F., González-Cabán, A., 2010. SINAMI: A Tool for the Economic
858 Evaluation of Forest Fire Management Programs in Mediterranean Ecosystems.
859 *International Journal of Wildland Fire* 19(7): 927–36.

860

861 Romero-Calcerrada, R., Barrio-Parra, F., Millington, J.D.A., Novillo, C.J., 2010. Spatial

862 modelling of socioeconomic data to understand patterns of human-caused wildfire
 863 ignition risk in the SW of Madrid (central Spain). *Ecological Modelling* 221, 34-45.
 864

865 Rothermel RC (1983) How to predict the spreads and intensity of forest and range fires.
 866 Gen. Tech. Rep. INT-143. Ogden, UT: U.S. Department of Agriculture, Forest Service,
 867 Intermountain Forest and Range Experiment Station. 161 p.
 868

869 Ruíz-Mirazo, J., Martínez-Fernández, J., Vega-García, C., 2012. Pastoral wildfires in
 870 the Mediterranean: Understanding their linkages to land cover patterns in managed
 871 landscapes. *Journal of Environmental Management* 98, 43-50.
 872

873 Scott, J.H., 2006. An analytical framework for quantifying wildland fire risk and fuel
 874 treatment benefit. In: Andrews, P.L., Butler, B.W. (Comps), *Fuels Management-How to*
 875 *Measure Success: Conference Proceedings*, March 28–30, Portland, OR. USDA Forest
 876 Service, Rocky Mountain Research Station Proceedings RMRS-P-41, pp. 169–184.
 877

878 Sebastián-López, A., Salvador-Civil, R., Gonzalo-Jimenez, J., SanMiguel-Ayanz, J.,
 879 2008. Integration of socio-economic and environmental variables for modelling long-
 880 term fire danger in southern Europe. *European Journal of Forest Research* 127, 149-163.
 881

882 Serra, P., Pons, X., Saurí, D., 2008. Land-cover and land-use change in a Mediterranean
 883 landscape: A spatial analysis of driving forces integrating biophysical and human
 884 factors. *Applied Geography* 28, 189-209.
 885

886 Taylor, S.W., Woolford, D.G., Dean, C.B., Martell, D.L., 2013. Wildfire Prediction to

887 Inform Fire Management: Statistical Science Challenges. *Statistical Science* 28(4): 586–
888 615.

889

890 Tuia, D., Lasaponara, R., Telesca, L., Kanevski, M., 2008a. Emergence of spatio-
891 temporal patterns in forest-fire sequences. *Physica a-Statistical Mechanics and Its*
892 *Applications* 387, 3271-3280.

893

894 Vasconcelos, M.J.P., Silva, S., Tome, M., Alvim, M., Pereira, J.M.C., 2001. Spatial
895 prediction of fire ignition probabilities: comparing logistic regression and neural
896 networks. *Photogrammetric Engineering and Remote Sensing* 67, 73-81.

897

898 Vega, J.A., Fernández, C., Jiménez, E., Ruíz, A.D., 2009b. Impacto de un escenario de
899 cambio climático sobre el peligro de incendios en Galicia. *Evidencias e impactos del*
900 *cambio climático en Galicia*. Xunta de Galicia. 581-206.

901

902 Vélez, R., 2000. Las quemas incontroladas como causa de incendios forestales.
903 *Cuadernos de la Sociedad Española de Ciencias Forestales* 9, 13-26.

904

905 Verde, J.C., Zêzere, J.L., 2010. Assessment and validation of wildfire susceptibility and
906 hazard in Portugal. *Nat. Hazards Earth Syst. Sci*, 10, 485-497.

907

908 Vicente, J., Crespo, F., 2012. A new wildland fire danger index for a Mediterranean
909 region and some validation aspects. *International Journal of Wildland Fire* 21, 1030-
910 1041.

911

912 Viedma, O., Angeler, D.G., Moreno, J.M., 2009. Landscape structural features control
 913 fire size in a Mediterranean forested area of central Spain. *International Journal of*
 914 *Wildland Fire* 18, 575-583.
 915
 916 Vilar del Hoyo, L., Woolford, D.G., Martell, D.L., Martín, M.P., 2010. A Model for
 917 Predicting Human-Caused Wildfire Occurrence in the Region of Madrid, Spain.
 918 *International Journal of Wildland Fire* 19(3): 325–337.
 919
 920 Westerling, A.L., Hidalgo, H.G., Cayan, D.R., Swetnam, T.W., 2006. Warming and
 921 Earlier Spring Increase Western U.S. Forest Wildfire Activity. *Science* 313(5789): 940–
 922 943.
 923
 924 Wittenberg, L., Malkinson, D., 2009. Spatio-temporal perspectives of forest fires
 925 regimes in a maturing Mediterranean mixed pine landscape. *European Journal of Forest*
 926 *Research* 128, 297-304.
 927
 928 Woolford, D.G., Bellhouse, D.R., Braun, W.J., Dean, C., Martell, D., Sun, J., 2011. A
 929 spatio-temporal model for people-caused forest fire occurrence in the Romeo Malette
 930 Forest. *Journal of Environmental Statistics*, 2(1): 1-26.
 931
 932 Wood, S.N., 2006. *Generalized Additive Models. An introduction with R.*
 933 *CRC/Chapman-Hall.*
 934
 935 Xunta de Galicia, 2014. Consellería de Medio Rural e de Mar. Avaliable at:
 936 http://www.medioruralemar.xunta.gal/nc/areas/forestal/ordenacion/man_comun/#inicio

[16 Jun 2014].

Figures and tables

Figure 1. Location of Galicia in Europe and average annual distribution of the number of wildfires in the EU by province (Source: Birot, 2009)

Figure 2. Spatial distribution of wildfires in the study area. Red points mark the centroid of the wildfire.

Figure 3. Estimated non-linear effects for weather and fuel model covariates, with their corresponding 80% and 95% pointwise credible intervals.

a Structured Spatial Effect – Posterior mode

b Posterior 95% probabilities of structured spatial effect

c Unstructured Spatial Effect – Posterior mode

d Posterior 95% probabilities of unstructured spatial effect

Figure 4. Structured and unstructured spatial effects on wildfires.

a Weather stations

b Distribution of average daily temperature

c Distribution of average relative humidity

d Distribution of rainfall

955

956 **Figure 5.** Area distribution in cases where weather covariate data were obtained from
957 weather stations.

958

Figure 6. Grid distribution in m² for the Rothermel fuel-model groups.

Variable	Description
Wildfires	"occurrence" (=1) and "non-occurrence" (=0) of wildfires
Altitude	Altitude recorded for each grid in m
Ta_mean	Average daily temperature recorded for each grid in °C
Hr	Daily average relative humidity for each grid in %
Pp	Daily rainfall recorded for each grid in l/m ²
Unallocated_model	Heterogeneous formations not assigned to any model for each grid in m ²
Pastures	Grassland formations for each grid in m ²
Shrub	Shrub formations for each grid in m ²
Litter under trees	Litter-under-tree formations for each grid in m ²

Wood_residuals

Forest formations with considerable quantities of dead woody residuals for each grid in m²

Cdconc

Municipalities in Galicia

Table 1. Variables generated from the baseline data

Model	-2*log-likelihood	df	AIC	GCV
GLM	15137.4	10	15137.4	0.1527
GLM plus spatial	14088.1	155.74	14399.6	0.0491
GAM	14992.3	25.35	15137.4	0.0527
STAR	11059.0	141.54	11342.0	0.0390

Table 2. Models fitted