Advances in Forest Fire Research

DOMINGOS XAVIER VIEGAS

EDITOR

2014

ANN multivariate analysis of factors that influence human-caused multiple fire starts

Sergi Costafreda-Aumedes[a], Cristina Vega-Garcia[b]

[a] *Agriculture and Forest Engineering Department, University of Lleida, Alcalde Rovira Roure 191, 25198, Lleida, Spain, E-mail: scaumedes@gmail.com*
[b] *Agriculture and Forest Engineering Department, University of Lleida, Alcalde Rovira Roure 191, 25198, Lleida, Spain, E-mail: cvega@eagrof.udl.es*

Abstract
Delays in the initial attack of new fire starts can happen locally when two or more fires burn simultaneously. The occurrence of multiple-fire-day situations may pose a real problem if suppression resources are limited, which almost always are. We analyzed multiple-fire-days in Galicia (Spain) from 2002 to 2005 with the goal of predicting these multiple fire situations by using Artificial Neural Networks. We carried out two types of analysis with our seasonally-structured data: to identify the relevant variables in the multiple versus single daily outcome (classification problem) and to predict the number of fires within the multiple-fire-days observations (prediction problem). The accuracy for the best Spring model was around 59-60% which located multiple occurrences in higher altitudes and public forest properties, near roads and recreation areas, with lower temperatures, lower quantity of pastureland and higher FFMC. Best classifications for the Summer period were around 60-61% and associated multiple fires to lower elevation areas, higher proportion of public and communal forests, near roads and higher drought indices. Predictions of actual number of fire occurrences in the Spring period reached 62% accuracy with a similar variables selection as the Spring classification model. Predictions for the Summer period lacked accuracy (44-50%) suggesting more complex patterns, probably due to mixed causes.

Keywords: *Artificial Neural Network, Daily Fire Prediction, Human-caused fires, Simultaneous fires*

1. Introduction

Short-term forest fire suppression performance depends on number and behavior of active fires (Haight and Fried 2007). Fire behavior is set by conditions in the fire environment that are difficult or impossible to control (fuels, weather, and topography), but numbers of ignitions in many countries, and particularly in the Mediterranean, are linked to human risk. Delays in the start of initial attack of new fires can happen locally when two or more forest fires are burning simultaneously, and the time required to extinguish a forest fire grows exponentially with detection and response time. For fire suppression resources sufficient to manage one fire, the simultaneous occurrence of two, three or more fires do create a challenge (Rachaniotis and Pappis 2006, IAFC and NFPA 2010). Forest fire managers must make crucial decisions every day on the amount and the type of fire suppression resources required and their allocation. Budgetary constraints under rising fire extinction costs (Liang *et al.* 2008, Calkin *et al.* 2014) often prevent maintaining enough suppression resources to potentially manage all possible active fires in all subdivisions of a region (i.e. Galicia, Alonso-Betanzos *et al.* 2003). When worse-case scenarios occur, and unlikely high-risk peaks couple with favorable conditions for burning in the fire environment, available fire suppression resources stretch beyond planned levels and are overloaded.
The prediction of these rare days with multiple fire starts, or multiple-fire-days (MFD), would be useful to fire managers (De Haan 2006). There are many studies which have focused on deployment of suppression resources (from Simard and Young 1978 to Calkin *et al.* 2014), but only Kirsch and Rideout (2005) and Haight and Fried (2007) tried to incorporate the effects of a number of possible simultaneous fires, but without estimating a probability or modeling this process. Previous studies

which attempted to predict MFD occurrence relied only on daily weather data or danger indices by using a Poisson distribution (Martell *et al.* 1987), binary logistic models (Andrews and Bradshaw 1997) or non-linear models (Garcia Diez *et al.* 1994, 1996, 1999). Garcia Diez *et al.* (1994, 1996, 1999) analyzed data between 1986 and 1993 in Galicia, concluding that fires (all causes) were primarily related to unfavorable weather conditions (past- and present-day weather). Their best MFD prediction had an accuracy varying between 52 and 72% over their study period.

Although MFD may be associated in some countries (Canada, USA) with dry lightning storms (Rorig *et al.* 2007), in most countries they are associated mainly to human activity (Omi 2005, Tanskanen and Venäläinen 2008, IAFC and NFPA 2010). Arsonists often set several ignition points to increase damage before firefighters arrive (Omi 2005, De Haan 2006), as do farmers in agrarian traditional rural activities (i.e. pastoral use of fire to regenerate rangelands in the Mediterranean, Ruiz-Mirazo *et al.* 2012). Martell *et al.* (1987), Garcia Diez *et al.* (1994, 1996, 1999) and Andrews and Bradshaw (1997) only considered weather variables related to the biophysical fire environment in their MFD occurrence analysis and did not include human risk factors.

There are already an important number of research papers on the general topic of fire occurrence prediction and the analysis of human risk in the fire literature (from Crosby 1954 to Rodrigues *et al.* 2014) and in Spain (i.e. Lozano *et al.* 2007, Chuvieco *et al.* 2009, Vasilakos *et al.* 2009, Vilar *et al.* 2010, Padilla and Vega-Garcia 2011, Vilar del Hoyo *et al.* 2011, Rodrigues *et al.* 2014). Fire occurrence studies have dealt with socioeconomic (i.e. Martínez *et al.* 2009) or geographic factors and biophysical variables (Lozano *et al.* 2007), but the occurrence of multiple-fire-day situations has been rarely approached.

Our aim was to explore these relationships between human risk factors, biophysical variables and multiple fires caused by people on a daily basis, with a view to aid fire management to anticipate and locate potential fire suppression overload events.

2. Methods

2.1. Study area

The study area is the Autonomous Community of Galicia, the Northwestern region of Spain (Figure 1). Galicia encompasses around 3 million ha (6 % of the Spanish territory) near the Atlantic Ocean (North and East boundaries). The humid Atlantic climate provides an average annual rainfall over 1,200 mm that favors the rapid growth of vegetation, which would be grasses and broadleaved tree species in natural conditions. The Galician vegetation has been highly modified through time. Population is distributed over 3,000 towns with a higher concentration on the coast and forests, mainly located in rural areas in the interior, that have long been highly fragmented by this human impact. Forest ownership is private (33 % of forests) either individually (64 % private forest area) or communally (36%). About 40% of forests have artificial origin, as many properties were reforested after the 1950's with *Pinus pinaster* and *Eucalyptus globulus* for their high economic interests for the pulp and plywood markets. Historical evidence attests the use of pastures by livestock for at least 6000 years (Kaal *et al.* 2011) and livestock production is currently the most important economic activity in the region.

Galicia is the Spanish administrative region with the highest number of human-caused forest fires (HCFF) (MAGRAMA 2013). Between 2001 and 2010, 86,036 HCFF have occurred in Galicia, a 44.1 % of all human-caused forest fires accounted for in the country in the same period, and the trend appears to be growing in this Atlantic region of the Iberian Peninsula (Carvalho *et al.* 2010). Fires take place basically into two well-defined time periods, end of winter and summer (Marey-Pérez *et al.* 2010), and are linked to socio-economic activities and the traditional pastoral use of fire (Marey-Pérez *et al.* 2010, Torre-Antón 2010). The high mean annual number of fires, and the spatiotemporal clustering of part of these fires in MFD in Galicia, led to its selection as study area.

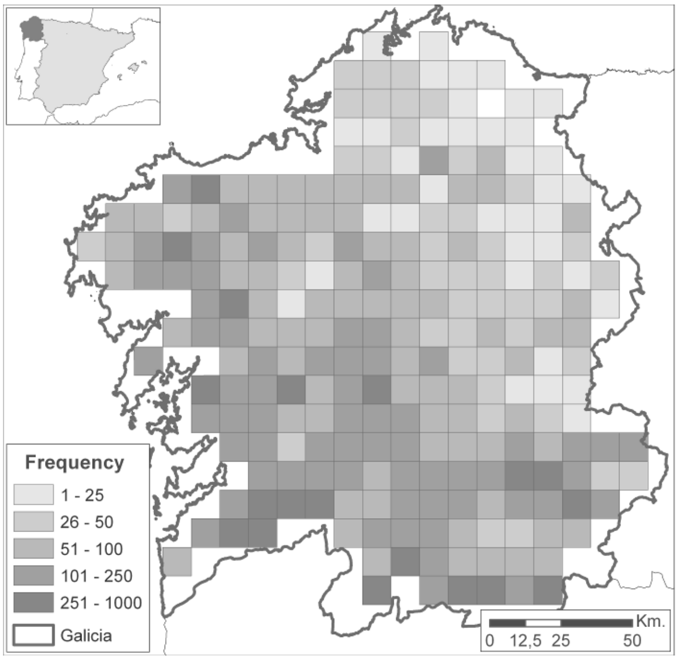

Figure 1. Galicia's location. This figure shows the human-caused forest fire frequency occurred during the period 2002 to 2005 in each of the 239 10x10 sq km UTM quadrates.

2.2. Study period

The data consisted on historical records of daily HCFF occurrences, daily weather data and geographic characteristics for Galicia in a 10x10 sq km UTM grid, for the period from 2002 to 2005. This period was restricted to four years due to data acquisition limitations (the spatially interpolated daily weather data available (Padilla and Vega-Garcia 2011)), but 4-5 years is the usual planning period for fire prevention in Spain.

2.3. Fire data

The Spanish fire history database (EGIF, 1983 – 2011) was provided by the Spanish Forest Service of the Ministry of Agriculture, Food and Environment (MAGRAMA). We extracted 28,442 HCFF records between 2002 and 2005 (both included) and summarized daily number of fires in a 10x10 sq km UTM grid for Galicia (19,612 observations). The grid consisted of 239 quadrates, after some irregular units in the coastal line were excluded in order to obtain a regular grid of equal-area (100 ± 1 km^2) quadrates. There was only one quadrate in which there was no fire in the study period. Though one fire occurrence per quadrate and day is quite common in Galicia (14,340 cases), MFD occurrences are rarer events (5,272 cases). In our study period, number of fires per day and quadrate ranged from 2 (3,354 observations) to 15 fires (just 1 observation).

The database was divided according to the seasonality of fires in Galicia (Marey-Pérez *et al*. 2010, Xunta de Galicia 2014). "Seasons" were defined by considering the daily frequency of the HCFF data, summarized by quadrate, and separating periods by minimum peaks of occurrence. Accordingly, the first period comprised from January 19th to May 11th (5,472 cases, 8,699 Spring fires), the second from May 12th to October 21st (13,735 cases, 19,248 Summer fires) and the third from October 22nd to January 18th (407 cases, 475 Winter fires). The third period had a very low number of records, so it was discarded for modeling.

We carried out two types of analysis with our data. Our first goal was the identification of variables that were relevant in the multiple (MFD) versus single (SFD) daily outcome. This classification analysis required coded data: SFD observations were coded as Y = (1, 0) (just one fire occurred in the day and quadrate) and MFD observations were coded Y = (0, 1) (more than one fire occurred in the quadrate and day). In order to have a balanced database for analysis (Vega-Garcia *et al.* 1996) we randomly selected the same number of cases for SFD and MFD records. This database contained 3,532 records in Spring and 6,900 in the Summer seasonal period (10,544 observations in total, annual data). The second goal was to predict quantitatively the number of HCFF within the MFD observations. We randomly selected the same number of observations for each class in order to have a balanced database for analysis. We limited the minimum number of observations per class to 100 because an increase of the daily fire frequency diminishes the number of cases and the modeling errors could increase. The maximum number of classes to predict was 6 (X = 1, 2, 3, 4, 5 and more than 5 HCFF per quadrate and day) when working with all the data (annual data) and five (X = 1, 2, 3, 4 and more than 4 HCFF per quadrate and day) when working with the data for each seasonal period. The seasonal MFD databases contained 820 and 1,055 observations for the Spring and Summer periods, respectively.

2.4. **Explanatory variables**

Our explanatory variables (Table 1) were the same as in the study of daily HCFF occurrence in Spain by Padilla and Vega-Garcia (2011). The cartographic information on these factors came from the Biodiversity database of the Spanish Ministry of Environment. Weather was provided by the Department of Geography of the University of Alcalá with permission of Meteológica SA (Madrid, Spain). The geographic factors included forest land ownership and infrastructures; physiography was described by elevation, slope and aspect variability; vegetation was represented by the percentage occupied by each Rothermel's fuel model (1972) in the quadrate. The weather data included daily records of rainfall, temperature, relative humidity, maximum dew point temperature, wind speed at 12 hours UTC, solar radiation, presence of snow and cloudiness. The Meteorological Fire Danger Index Processor (MFDIP) described in Camia *et al.* (1998) was used to compute several fire danger rating indices widely used.

2.5. **Methodological approach. Cascade correlation ANN**

Artificial Neural Networks (ANN) models are a reliable alternative to traditional statistical methods, especially when dealing with large databases, non-linear systems (Scrinzi *et al.* 2007), not normally distributed variables (Hilbert and Ostendorf 2001) or highly correlated variables (Vega-Garcia and Chuvieco 2006). For a given database representing a set of historical events these algorithms are capable of identifying and fitting very complex non-linear patterns by iterative adjustment (Kuplich 2006). These models have been applied before to the study of forest fires (e.g. Vega-Garcia *et al.* 1996, Vasconcelos *et al.* 2001, Alonso-Betanzos *et al.* 2003, Vega-Garcia and Chuvieco 2006, Vasilakos *et al*. 2009) reaching good accuracies and good generalization capability, so ANN models were selected for this work in an attempt to gauge their potential in modeling MFD events.

As losing generalization capacity by overtraining a net is a common problem, it is customary to apply an *early stopping* procedure, based on the separation of the database in two subsets: one for training or "learning", and part (test set) used to control the evolution of error committed in this independent set, and stop training if the error starts to increase (Hasenauer *et al*. 2001, Jutras *et al*. 2009). Accordingly, before construction of the models the database was randomly split in two subsets: 70 % of the cases were used for the training process, and the remaining 30 % were used to test the progression of the iterative learning processes and to prevent overtraining of the models that might limit their generalization capacity. No independent validation dataset was segregated (100% cases were used for validation), as the models were designed to be explicative and not to predict future occurrences.

Table 1. Acronyms and abbreviations of each independent variable

Acronyms and Abbreviations	Variable	Acronyms and Abbreviations	Variable
Geographical variables		Geographical variables	
Human Risk Factors		Physiography	
Forest Land Tenure		DEM_MN	Elevation mean
FOREST_PV	Private	DEM_SD	Standard deviation of the elevation
FOREST_PB	Public	SLOPE_MN	Slope mean
FOREST_VC	Communal	SLOPE_SD	Standard deviation of the slope
FOREST_PR	Protected	ASPECT_VAR	Number of aspect classes
Infrastructure		Meteorological variables	
TOWN_DI	Distance to towns		
RECRE_DI	Distance to recreational areas	Raw weather variables	
RESER_DI	Distance to reservoirs	P24	Daily rainfall
ROAD_DI	Distance to roads	T_MAX	Maximum temperature
DENS_PO	Density of population	RH_MIN	Minimum relative humidity
		DewTmpMax	Maximum dew point temperature
Vegetation		WIND	Wind speed at 1200 hours UTC
Area occupied by each	Rothermel's fuel type (%)	SOLAR_R	Solar radiation
FUEL_01	Fuel type 1	CLOUD	Cloudiness
FUEL_02	Fuel type 2		
FUEL_03	Fuel type 3	Fire danger rating indices	
FUEL_04	Fuel type 4	FWI	Canadian Fire Weather Index
FUEL_05	Fuel type 5	FFMC	Canadian Fine Fuel Moisture Code
FUEL_06	Fuel type 6	DMC	Canadian Duff Moisture Code
FUEL_07	Fuel type 7	DC	Canadian Drought Code
FUEL_08	Fuel type 8	PROB	Portuguese Drought Index
FUEL_09	Fuel type 9	ISI	Initial Spread Index
FUEL_00	No vegetation		
FUEL_TY	Number of different fuel types		

The backpropagation algorithm (Rumelhart *et al*. 1986) associated to the multilayer perceptron (Hecht-Nielsen 2005), with one hidden layer and a sigmoid activation function, has been the algorithm most often used for predicting fire occurrence (Vega-Garcia *et al*. 1996, Vasilakos *et al*. 2009), but other techniques such as genetic algorithms (Vasconcelos *et al*. 2001), Levenberg-Marquardt models (Alonso-Betanzos *et al*. 2003), perceptron neural networks (Bisquert *et al*. 2012) or cascade-correlation networks (Vega-Garcia and Chuvieco 2006) have also been used for modeling forest fire occurrence.

The cascade-correlation model created by Fahlman and Lebiere (1990) was used in this work, following the procedure described in Alcázar *et al.* (2008) and Vega-Garcia and Chuvieco (2006). Initial architecture was set to two layers (input/output), and no hidden units. The number of hidden nodes was optimized during the training phase, as candidate hidden units were tested and added (or not), creating an optimal multi-layer structure (Fahlman and Lebiere 1990) at the end of the learning phase. We controlled final size to make sure number of weights was within the order of magnitude that our data would allow, and favored simple architectures. Training was based on an adaptative gradient learning rule (Bridle 1990, Fahlman and Lebiere 1990), a fast variant of the general algorithm of back-propagation (Werbos 1994) that featured a weight decay factor to reduce complexity of the models (NeuralWare 2009). Several models were built and tested starting training with different random weights, and with different random distributions of individual observations among the three datasets (training, test and validation).

Classification performance for the two cases (SFD and MFD) in the three datasets (training, test and validation) was evaluated by using the total percentage correctly classified, or Accuracy, and the confusion matrix provided by Predict® 3.24 software (NeuralWare 2009). The predictive capacity of the number of daily forest fires (X = 1, 2, 3, 4, 5, 6) was evaluated by using the linear correlation (r) between the observed and the estimated fire occurrences in three data subsets (training, testing and validation). In the selection of the best models, other diagnostics were analyzed, such as relative entropy of the network (the lower, the better the model was), internal correlation (the higher, the better) (NeuralWare 2009), and architecture complexity (we favored models with low number of input variables, processing elements and layers as stated before).

We sought balanced accuracies for the three datasets. For the best model, a sensitivity analysis was applied to identify relevant variables by computing a matrix of partial derivatives of the output variables with respect to each of the input variables (NeuralWare 2009). High values of this sensitivity measure indicated that slight variations of an input variable produced considerable changes in the classification of SFD or MFD, and vice versa. A positive sign in the sensitivity analysis would generally indicate a direct relationship between dependent and independent variable; a negative sign would indicate an inverse relationship. This is a standard diagnostic procedure commonly used to gain insight into a multilayer neural network solution (Alonso-Betanzos *et al.* 2003, Alcázar *et al.* 2008).

3. Results

Cascade-correlation binary classification accuracies were around 60% with a 7-8-2 ANN structure, for the annual data. This model suggested that MFDs occurred in areas with lower elevation (DEM_MN, average MDF value: -0.200) and lower temperatures (T_MAX, -0.107), higher FFMC (1.321) and higher DC (0.125).

Table 2. Accuracy and sensitivity analysis of the Spring cascade-correlation binary classification model (CL_SP). Network structure: 6-2-2

	Accuracy	Relative Entropy		FOREST_PB	ROAD_DI	RECRE_DI	FUEL_02	DEM_MN	FFMC
SFD	0.653	0.639	Av. SFD	-0.411	0.441	0.268	0.413	-0.392	-1.469
MFD	0.576	0.676	Av. MFD	0.411	-0.441	-0.268	-0.413	0.392	1.469
Train	0.615	0.657							
SFD	0.625	0.649	Av. Sq.	0.204	2133	0.219	0.237	0.201	3.938
MFD	0.582	0.698	Variance	0.036	1939	0.148	0.066	0.047	1.779
Test	0.603	0.674							
SFD	0.645	0.642							
MFD	0.578	0.683							
Validation	0.611	0.662							

SFD, Single-Fire-Days; MFD, Multiple-Fire-Days; Av. SFD, Average SFD value per variable; Av. MFD, Average MFD value per variable; Av. sq. Average square of each variable

Accuracy for the best Spring model (CL_SP) was also around 59-60% (Table 2), but DEM_MN had a positive response (contrariwise to the annual binary classification model). T_MAX was included in some models evaluated, with negative sign. MFD occurred more frequently in higher altitudes, with higher percentages of public forests (FOREST_PB), where temperatures are lower and near roads

(ROAD_DI) and recreation areas (RECRE_DI), with lower quantity of pasture FUEL_02 and higher FFMC.

Best classifications for the second seasonal period (Summer, CL_SU) were located again around 60-61%, similar to the first period. T_MAX was included only in few models, but in those cases, the sign of T_MAX was positive. MFD (Table 3) were associated to lower elevation areas with higher proportion of public and communal forests (FOREST_PB and FOREST_VC), near roads, and higher drought indices (DMC and DC).

Table 3. Accuracy and sensitivity analysis of the Summer cascade-correlation binary classification model (CL_SU). Network structure: 7-4-2

	Accuracy	Relative Entropy		FOREST_PB	FOREST_VC	ROAD_DI	DEM_MN	DMC	DC
SFD	0.595	0.663	Av. SFD	-0.056	-0.118	0.672	0.220	-0.226	-0.182
MFD	0.607	0.662	Av. MFD	0.056	0.118	-0.672	-0.220	0.226	0.182
Train	0.601	0.663							
SFD	0.609	0.659	Av. Sq.	0.004	0.097	0.575	0.425	0.067	0.227
MFD	0.614	0.669	Variance	0.001	0.083	0.124	0.377	0.016	0.194
Test	0.611	0.664							
SFD	0.599	0.662							
MFD	0.609	0.664							
Validation	0.604	0.663							

SFD, Single-Fire-Days; MFD, Multiple-Fire-Days; Av. SFD, Average SFD value per variable; Av. MFD, Average MFD value per variable; Av. sq. Average square of each variable

3.1. Numerical models

No model could be accepted to adequately discriminate the exact number of human-caused forest fires per day and quadrate for the annual data, so we focused our efforts on the seasonal models.

Table 4. Accuracy and sensitivity analysis of the Spring cascade-correlation predictive model (PR_SP). Network structure: 11-14-1

PREDICT	R	Av. Abs.	Max. Abs.	RMS	Accuracy (20%)	Conf. Interval (95%)		
Train	0.609	0.927	2.950	1.109	0.471	2.161		
Test	0.597	0.958	2.957	1.162	0.462	2.273		
Validation	0.606	0.936	2.957	1.125	0.468	2.191		
	FOREST_PB	FOREST_PV	DENS_POB	FUEL_06	SLOPE_MN	SLOPE_SD	PROB	DMC
Average	-0.273	0.324	-0.988	1.531	0.187	-0.198	0.276	0.720
Average Square	0.123	0.232	3.857	43.264	1.332	0.142	0.151	0.844
Variance	0.049	0.127	2.884	40.969	1.299	0.103	0.075	0.326

SFD, Single-Fire-Days; MFD, Multiple-Fire-Days; Av. Abs., Average Absolute Error; Max. Abs., Maximum Absolute Error

Predictions of fire occurrences (1 to 5 values) in the Spring seasonal period (PR_SP) reached 61% accuracy (Table 4). The structure of the net was relatively complex (high number of variables in the hidden layer, structure 11-14-1). The average error was approximately 1 fire, with an absolute error of 3. In this case, higher HCFF occurred in quadrates with higher percentage of private forest (FOREST_PV), but lower public forest (FOREST_PB), lower population density, higher shrub fuel (FUEL_06) percentage located in high slopes, and higher PROB and DMC (drought indices).

Results in the Summer period (PR_SU) were somewhat worse (44-50% in the best model, Table 5). Drought indices (DC, DMC, etc.) were the most often selected variables. In the best model, the quantity of input variables was high, but the net had only 3 nodes in the hidden layer (24-3-1). A large quantity of anthropogenic-related variables was included in this net model, locating fires mainly near roads and recreational areas, but far from towns and reservoirs. This model indicated that HCFF were more numerous in locations near roads and recreational areas, higher proportion of public and protected forests, lesser percentage of private and communal forests, lesser altitude and slope variation, and higher danger and drought indices.

Table 5. Accuracy and confusion matrix of the Summer cascade-correlation predictive model (PR_SU). Network structure: 24-3-1

PREDICT	R	Av. Abs.	Max. Abs.	RMS	Accuracy (20%)	Conf. Interval (95%)	Records
Train	0.437	1.102	3.214	1.288	0.358	2.510	738
Test	0.499	0.986	2.910	1.187	0.423	2.317	317
Validation	0.456	1.067	3.214	1.259	0.377	2.450	1055

Av. Abs., Average Absolute Error; Max. Abs., Maximum Absolute Error

4. Discussion

Multiple fire days in quadrates in Galicia could be correctly separated from the more usual occurrence of single fire days in seasonal time spans using Cascade-Correlation models. Variables and trends were similar in most of the models evaluated, which suggested robustness in the model building processes and relevance of these variables in the final solution. Binary classification seasonal models (X = 1 or more than 1 human-caused forest fire per day and quadrate) have higher accuracy than predictive models (X = 1, 2, 3, 4 or more than 4 HCFF per day and quadrate) especially for the Spring "season". This better accuracy suggests that Spring fires caused by humans are probably more homogeneous and related to a smaller number of causes mainly linked to vegetation management, like agricultural residues burning or shrubs removal (Chas-Amil *et al.* 2010a). On the contrary, the lower accuracy and the higher number of variables of the Summer models seem to reveal that human-caused forest fires in summer are more heterogeneous in cause (i.e. agricultural burning, shrubs removal, arsonists, negligence, machinery), making especially difficult trying to predict the number of fires per quadrate and day (Morillo 2013, personal communication).

Our models show that the probability of MFD was influenced by structural or geographic variables -population, forest ownership, vegetation and topography- and by temporally changing daily variables -weather variables and fire danger assessment indices- during the study period. While accuracy was similar in all the models (around 60%), some variables were different or exhibited opposite trends in them. These findings are in agreement with known causes of fires and their temporal distribution, as described previously (Chas-Amil *et al.* 2010a). Our results indicate that fire activity can vary considerably between different locations in a given day in Galicia in agree to Marey-Perez *et al.* (2010), as in other work elsewhere by i.e. Boychuck and Martell (1988) in Canada or Vasilakos *et al.* (2009) in Lesbos Island (Greece).

Regarding weather variables, the maximum temperature factor influence varied with season, but fire danger indices were positively related to higher fire incidence in all seasons. Spring and Summer MFD were more likely on windy days with lower relative humidity, which may be translated as days with higher danger indices (FFMC, DMC, DC or PROB) in agreement with Garcia Diez *et al.* (1994, 1996, 1999) and Vasilakos *et al.* (2009). The Galician weather is humid and rainy most of the year except in Summer, but there are inland areas where a geographic gradation towards the Mediterranean climate creates drier weather conditions (and higher danger indices) than in the Atlantic region. In Galicia, danger indices may be high when dry weather conditions prevail, such as in summertime with the higher maximum temperatures and fuel moisture deficit, but also with low daily relative humidity and atmospheric stability in winter (colder days) (Garcia Diez *et al.* 1999) which explains the opposite seasonal temperature pattern found in different models. Spells of low relative humidity (RH) in late winter – early spring produce vegetation water stress during these periods and facilitates fuel ignition, which is why shrublands burning takes place mainly in late winter - early spring, but it is also carried out during the summer, when is masked by other agricultural burning activities and other arson or negligence fires (Chas-Amil *et al.* 2010b, Morillo 2013, personal communication).

The selection of the structural/geographic variables is consistent with the descriptions of the fire problem in Galicia by Chas-Amil *et al.* (2010a, 2010b) and also helps to understand this pronounced seasonality. Altitude is the main distinctive geographic factor. Spring MDFs are more usual in higher altitudes (generally inland); meanwhile Summer MDFs mainly occur in lower altitudes (generally coastal).

According to the Spanish forest fire database, about 75 % of human-caused forest fires in Galicia are related to the rural use of fire for vegetation management (Torre-Antón 2010). Fire use is common in Southern rural areas inland, with higher altitudes and slopes, where most wildfires take place (Chas-Amil *et al.* 2010a, 2010b). Rural population concentrates in valley areas with high livestock productivity, mainly cattle, which feed on grasses. Livestock management requires grasses and the cheapest and traditional way to eliminate shrubs, mainly gorse and heather (Díaz-Vizcaino 2005), is the use of fire (Kaal *et al.* 2011). The second most common fire cause in Galicia is agricultural residues burning. Agriculture is also an important economic activity in Galician rural areas and stubble burning is quite common (MAGRAMA 2013).

Lower altitudes are located in Western and Northern Galicia (Fig 1., at the boundary with the Atlantic Ocean). According to Chas-Amil *et al.* (2010b), in Western Galicia, human-caused forest fires have more diverse motivations. The Southwest is dominated by arsonists; meanwhile the Northwest is dominated by unspecified motivations as revenge, timber prices, land use changes or resentments. Other near township's fires are related to agricultural and livestock activities. As in Vasilakos *et al.* (2009), our models related summer MDFs to lower altitudes in which human settlements (towns, recreational areas, etc.) are more usually placed and recreation and agriculture are the most important economic activities (Chas-Amil 2007). These activities are related to forest fires especially in the boundary of human activities-forest interfaces (Martinez *et al*. 2010). In these Galician regions forest are not valued by the population (Marey-Pérez *et al.* 2010, Torre-Antón 2010), which increases the risk of fires. Public forests and Communal forests were selected in our best Summer binary and predictive models, also in agreement to the conclusions of Marey-Perez *et al.* (2010).

Useful as these explanatory models may be, we are driven to consider that complex spatiotemporal patterns exist. In order to better to explore these relationships between human risk factors, biophysical variables and multiple fires caused by people on a daily basis, with a view to aid fire management to anticipate and locate potential fire suppression overload events, causality will have to be specifically incorporated in future models. However, cause records are currently incomplete, as full investigations are precluded by the high daily fire loads.

5. Acknowledgements

he authors gratefully acknowledge the fire occurrence historical data from the National Forest Fire Statistics database (EGIF), Ministry of Environment and Rural and Marine Affairs (MAGRAMA) and weather data from the Department of Geography of the University of Alcalá and Meteológica SA (Madrid, Spain). Arsenio Morillo (Xunta de Galicia) provided helpful advice and on-site expertise.

6. References

Alcázar J, Palau A, Vega-García C (2008) A neural net model for environmental flow estimation at the Ebro River Basin, Spain. *Journal of Hydrology* **349**, 44-55.

Alonso-Betanzos A, Fontenla-Romero O, Guijarro-Berdiñas B, Hernández-Pereira E, Paz Andrade MI, Andrade P, Jiménez E, Legido Soto JL, Carballas T (2003). An intelligent system for forest fire risk prediction and fire fighting management in Galicia. *Expert Systems with Applications* **25**, 545-554.

Andrews PL, Bradshaw LS (1997) The FIRES computer program. General Technical Report - US Department of Agriculture, Forest Service, Issue INT-GTR-365

Boychuk D, Martell DL (1988) A Markov Chain Model for Evaluating Seasonal Forest Fire Fighter Requirements. *Forest Science* **34**, 647-661.

Bridle JS (1990) Alpha-nets: A recurrent 'neural' network architecture with a hidden Markov model interpretation. *Speech Communication* **9**, 83-92.

Camia A, Bovio G, Gottero F (1998) Algorithms to compute the Meteorological Danger Indices included in MFDIP. Megafires Project, Final Report.

Carvalho A, Flannigan MD, Logan KA, Gowman LM, Miranda AI, Borrego C (2010) The impact of spatial resolution on area burned and fire occurrence projections in Portugal under climate change. *Climatic Change* **98**, 177-197.

Chas-Amil ML (2007) Forest fires in Galicia (Spain): Threats and challenges for the future. *Journal of Forest Economics* **13**, 1-5.

Chas-Amil ML, Touza J, Prestemon JP (2010a) Spatial distribution of human-caused forest fires in Galicia (NW Spain). *Ecology and the Environment* **137**, 247-258.

Chas-Amil ML, Touza J, Prestemon JP, McClean Colin J (2010b) Sociorvonomic analysis of forest fires causes and motivations: A case study of Galicia. VI International Conference on Forest Fire Research. In: Viegas DX (Ed), Coimbra, Portugal.

Chuvieco E, González I, Verdú F, Aguado I, Yebra M (2009) Prediction of fire occurrence from live fuel moisture content measurements in a Mediterranean ecosystem. *International Journal of Wildland Fire* **18**, 430-441.

Crosby JS (1954) Probability of fire occurrence can be predicted. *USDA, Forest Service, Central States Forest Experiment Station, Technical Paper* **143**, 14-14.

De Haan JD (2006) Kirk's Fire Investigation. Prentice Hall, ISBN 9780131719224

Donovan GH, Rideout DB (2003) A reformulation of the cost plus net value change (C+NV) model of wildfire economics. *Forest Science* **49**, 318-323.

Falham SE, Lebiere C (1990) The Cascade-Correlation Learning Architecture. Advances in Neural Information Processing Systems 2. Morgan Kaufmann, San Francisco, CA, USA. 13 pp

Garcia Diez EL, Rivas Soriano L, de Pablo F, Garcia Diez A (1994) An objective model for the daily outbreak of forest fires based on meteorological considerations. *Journal of Applied Meteorology* **33**, 519-526.

Garcia Diez EL, Rivas Soriano L, Garcia Diez A (1996) Medium-range forecasting for the number of daily forest fires. *Journal of Applied Meteorology* **35**, 725-732.

Garcia Diez EL, Rivas Soriano L, de Pablo F, Garcia Diez A (1999) Prediction of the daily number of forest fires. *International Journal of Wildland Fire* **9**, 207-211.

Haight RG, Fried JS (2007) Deploying Wildland Fire Suppression Resources with a Scenario Based Standard Response Model. *INFOR* **45**, 31-39.

Hassenauer H, Merkl D, Weingartner M (2001) Estimating tree mortality of Norway spruce stands with neural networks. *Advances in Environmental Research* **5**, 405-414.

Hilbert DW, Ostendorf B (2001). The utility of artificial neural networks for modelling the distribution of vegetation in past, present and future climates. *Ecological Modelling* **146**, 311–327.

IAFC, NFPA (Eds) (2010) Fire Officer: Principles and Practice. Jones and Bartlett Publishers, Sudbury, MA, USA

Islam KMS, Martell DL (1998) Performance of initial attack airtanker systems with interacting bases and variable initial attack ranges. *Canadian Journal of Forest Research* **28**, 1448-1455.

Jutras P, Prasher SO, Mehuys GR (2009) Prediction of street tree morphological parameters using artificial neural networks, Computers and Electronics in Agriculture 67, 9-17.

Kaal J, Carrión Marco Y, Asouti E, Martín Seijo M, Martínez Cortijas A, Costas-Casáis M, Criado Boado F (2011) Long-term deforestation in NW Spain: linking the Holocene fire history to vegetation chance and human activities. *Quaternary Science Reviews* **30**, 161-175.

Kirsch AG, Rideout DB (2005) Optimizing initial attack effectiveness by using performance measures. System Analysis in Forest Resources: Proceedings of the 2003 Symposium In: (Eds M Bevers, TM Barrett). pp. 183-187 (Portland, OR)

Kuplich TM (2006) Classifying regenerating forest stages in Amazonia using remotely sensed images and a neural network. *Forest Ecology and Management* **234**, 1-9.

Liang J, Calkin DE, Gebert KM, Venn TJ, Silverstein RP (2008) Factors influencing large wildland fire suppression expenditures. *International Journal of Wildland Fire* **17**, 650-659.

Lozano FJ, Suárez-Seoane S, de Luis E (2007) Assessment of several spectral indices derived from multi-temporal Landsat data for fire occurrence probability modelling. *Remote Sensing of Environment* **107**, 533-544

Marey-Pérez MF, Gómez-Vázquez I, Díaz-Varela E (2010) Different approaches to the social vision of communal land Management: the case of Galicia. *Spanish Journal of Forest Research* **8**, 848-863.

Martell DL, Otukol S, Stocks BJ (1987) A logistic model for predicting daily people-caused forest fire occurrence in Ontario. *Canadian Journal of Forest Research* **17**, 394-401.

Martínez J, Vega-Garcia C, Chuvieco E (2009) Human-caused wildfire risk rating for prevention planning in Spain. *Journal of Environmental Management* **90**, 1241-1252.

MAGRAMA (2013) Forest Fires in Spain 2012, Area of Defense Against Forest Fires. Madrid.

Morillo A (2013) Coordinator of Prevention and Forest Valoration Area of Consellería do Medio Rurale do Mar, Xunta de Galicia, Spain.

NeuralWare (2009) NeuralWorks Predict®. The complete solution for neural data modeling. NeuralWare, Pittsburg, PA, USA

Omi PH (2005) Forest fires: a reference handbook. Contemporary world issues. Santa Barbara, California, USA

Padilla M, Vega-Garcia C (2011) 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. *International Journal of Wildland Fire* **20**, 46-58.

Rachaniotis NP, Pappis CP (2006) Scheduling fire-fighting tasks using the concept of "deteriorating jobs". *Canadian Journal of Forest Research* **36**, 652-658.

Rorig ML, McKay SJ, Ferguson SA, Werth P (2007) Model-Generated Predictions of Dry Thunderstorm Potential. *Journal of Applied Meteorology and Climatology* **46**, 605-614.

Rothermel RC (1972) A mathematical model for predicting fire spread in wildland fuels, Res. Pap. INT-115. Ogden, UT: U.S. Department of Agriculture, Intermountain Forest and Range Experiment Station

Rumelhart DE, Hinton HE, Williams RJ (1986) Learning integral representations by error propagation. In: Parallel distributed processing: explorations in the microstructure of cognition. Foundations (Eds DE Rumelhart, JL McClelland). MIT Press Vol. 1. pp. 318-362.

Scrinzi G, Marzullo L, Galvagni D (2007) Development of a neural network model to update forest distribution data for managed alpine stands. *Ecological Modelling* **206**, 331-346.

Simard AJ, Young A (1978) AIRPRO - an air tanker productivity computer simulation model - application. Ontario Information Report FF-X-69. (Forest Fire Research Institute, Canadian Forest Service: Ottawa)

Tanskanen H, Venäläinen A (2008) The relationship between fire activity and fire weather indices at different stages of the growing season in Finland. *Boreal Environment Research* **13**, 285-302.

Torre-Antón M (2010) La situación actual de los incendios forestales en España. Conferencias y Ponencias del 5º Congreso Forestal Español, *Cuadernos de la Sociedad Española de Ciencias Forestales* **31**, 179-195.

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

Vasilakos C, Kalabokidis K, Hatzopoulos J, Matsinnos I (2009) Identifying wildland fire ignition factors through sensitivity analysis of a neural network. *Natural Hazards* **50**, 125-143.

Vega-Garcia C, Chuvieco E (2006) Applying local measures of spatial heterogeneity to Landsat-TM images for predicting wildfire occurrence in Mediterranean landscapes. *Landscape Ecology* **21**, 595-605.

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

Vilar L, Woolford DG, Martell DL, Martín MP (2010) A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. *International Journal of Wildland Fire* **19**, 1-13.

Vilar del Hoyo L, Martín MP, Martínez-Vega FJ (2011) Logistic regression models for human-caused wildfire risk estimation: Analysing the effect of the spatial accuracy in fire occurrence data. *European Journal of Forest Research* **130**, 983-996.

Werbos, PJ (1994) The roots of backpropagation: from ordered derivatives to neural networks and political forecasting, John Wiley & Sons, Inc, New York, USA, 319 pp.

Xunta de Galicia (2014) Plan prevención e defensa contra os incendios forestais de Galicia (PLADIGA) Memoria 2014. Consellería do medio rural, Dirección Xeral de Montes.