An Acad Bras Cienc (2023) 95(2): e20201039 DOI 10.1590/0001-3765202320201039
Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*
Printed ISSN 0001-3765 I Online ISSN 1678-2690
www.scielo.br/aabc | www.fb.com/aabcjournal

FORESTRY SCIENCE

Use of machine learning as a tool for determining fire management units in the brazilian atlantic forest

RONIE S. JUVANHOL, NILTON CESAR FIEDLER, ALEXANDRE R. DOS SANTOS, TELMA M.O. PELUZIO, WELLINGTON B. DA SILVA, CHRISTIANO JORGE G. PINHEIRO & HELBECY CRISTINO P. DE SOUSA

Abstract: Geoprocessing techniques are generally applied in natural disaster risk management due to their ability to integrate and visualize different sets of geographic data. The objective of this study was to evaluate the capacity of classification and regression tree (CART) to assess fire risk. MCD45A1 product of the burnt area, relative to a 16-year period (2000-2015) was used to obtain a fire occurrence map, from center points of the raster, using a kernel density approach. The resulting map was then used as a response variable for CART analysis with fire influence variables used as predictors. A total of 12 predictors were determined from several databases, including environmental, physical, and socioeconomic aspects. Rules generated by the regression process allowed to of define different risk levels, expressed in 35 management units, and used to produce a fire prediction map. Results of the regression process ($r = 0.94$ and $r^2 = 0.88$) demonstrate the capability of the CART algorithm in highlighting hierarchical relationships among predictors, while the model's easy interpretability provides a solid basis for decision making. This methodology can be expanded in other environmental risk analysis studies and applied to any area of the globe on a regional scale.

Key words: non-parametric statistics, kernel density, cart algorithm, decision rules, fire prediction map.

INTRODUCTION

Over the past few decades, forest fires in Brazil have received greater attention due to their wide range of ecological, economic, social and political impacts, although statistics of fire occurrence and their effects are still incipient.

Fire plays an important role in creating and maintaining landscape structure, composition, function and ecological integrity, and may influence the rates and processes of ecological succession (Covington & Moore 1994, Morgan et al. 2001). Fire impact on a local, regional and global scale was revised in Stolle and Lambin (2003) and Lentile et al. (2006). On local scale, fire can stimulate microbial processes and alter structure and composition of soils as well vegetation (Lentile et al. 2006). At regional and global scales, combustion of forest and grassland vegetation releases large volumes of active gases, pyrogenic aerosols and other compounds that significantly influence Earth's radiative budget and chemistry atmospheric (Andreae & Merlet 2001), affecting air quality (Hardy et al. 2001) and raising concerns about risks to human health (Stefanidou et al. 2008).

Spatial distribution of forest fires and their main promoting factors is mainly understood by the analysis of historical occurrence data

(Syphard et al. 2008). The importance of anthropogenic factors in regulation of fire events, in addition to climate, vegetation and topographic ones, makes fire prediction highly challenging (Perry 1998). Thus, development and use of fire prediction models can assist in forest management active and preventive decision-making (González et al. 2006).

A wide range of techniques has been used to model fire risk. More complex models of fire require spatial information that is provided by remote sensing and Geographic Information Systems (GIS) (Bonazountas et al. 2005). The use of data from multiple sources, entails the presence of local variation and multivariate relationships among predicting variables, what demands flexible and consistent models. Nevertheless, commonly used models suggest an a priori setting of modeling parameters. This preliminary setup is often based on knowledge of fire experts (Caetano et al. 2004, Chuvieco & Congalton 1989, Eugenio et al. 2016, Semeraro et al. 2016, Vadrevu et al. 2010) or in regression analysis, where coefficients represent the weights of considered predictive variables.

Statistical models proposed in previous studies range from multiple linear regression (Oliveira et al. 2012) to logistic regression (Bem et al. 2018, Catry et al. 2009, Kalabokidis et al. 2007, Martínez et al. 2009, Syphard et al. 2008). According to Amatulli et al. (2006), multiple linear regression is limited once it does not consider discrete variables in the model, such as land use and type of fuel, which is very important in modeling fire risk. Meanwhile, in logistic regression models, the output target variable is a simple "occurs / does not occur" answer. This limitation restricts its use when the target variable is continuous, such as fire density (fire events per unit area). Furthermore, according to Martínez-Fernández et al. (2013), the regression coefficients obtained by these statistical techniques are applied to the entire study area, without considering the spatial variation of each variable in the regression process. This assumption, also known as stationarity, is ignored at the time of modeling.

In order to overcome these disadvantages, some authors have used neural networks to predict fire risk, with satisfactory results (Chuvieco 1999, Vasconcelos et al. 2001, Vega-Garcia et al. 1996). Although neural networks computes accurate fire risk maps, it does not provide information about the variables degree of importance in the predictive model.

As presented, different techniques have been tested and developed, but each with its shortcomings in the prediction of forest fires. Therefore, the applicable models must be more robust to be used as efficient operational tools in the management of forest fires. Thus, they must be able not only to deal with different data sources, but also to provide a more detailed result than a single output risk map.

Against mentioned gaps of current fire prediction systems, machine learning algorithms offer an interesting approach to treat the problem. Some algorithms have been applied to predict fire in the literature, such as Random Forest (Arpaci et al. 2014, dos Santos et al. 2020, Wu et al. 2014), MaxEnt (Arpaci et al. 2014, Fonseca et al. 2016, Martín et al. 2019), and Boosted Regression Trees (Argañaraz et al. 2015). However, these approaches do not have the fascination of creating a single tree structure with easy interpretation for managers, which is offered by Classification and Regression Trees algorithm (CART).

Proposed by Breiman et al. (1984), CART decision tree can process continuous and categorical attributes through a binary recursive procedure that constructs an ideal tree. Regarded as one of the ten best second data mining algorithms Wu & Kumar (2009), classification tree predicts a likelihood of association for categorical response variables, while regression tree provides average values for continuous response variables in interval or scale of reason (Michaelsen et al. 1994), which is useful for fire prediction.

This study falls within spatial modeling and analysis using machine learning techniques to evaluate fire prediction on regional scale; a field in which the potential of this statistical approach has not yet been well explored. The proposed technique aims to provide understandable outputs, in the form of decision rules, enabling to predict average risk values for each grid cell and set the fire management units in Espírito Santo State.

MATERIALS AND METHODS

Area of study

The study area is represented by Espírito Santo State, located in Southeast region of Brazil (Figure 1), with an area of 46,052.64 km². It is located between 17°53'29'' and 21°18'03'' of latitude S and 39° 41'18'' and 41°52'45'' of longitude W. It borders Atlantic Ocean to the East, Bahia State to North, Minas Gerais State to West and Rio de Janeiro State to South. Due to its geographical location and geomorphology, the state presents four types of climate according to Köppen classification: Cwb, subtropical climate of altitude with dry winter and mild summers, found in the state mountainous region; Cwa, subtropical climate of dry winter and hot summer, found

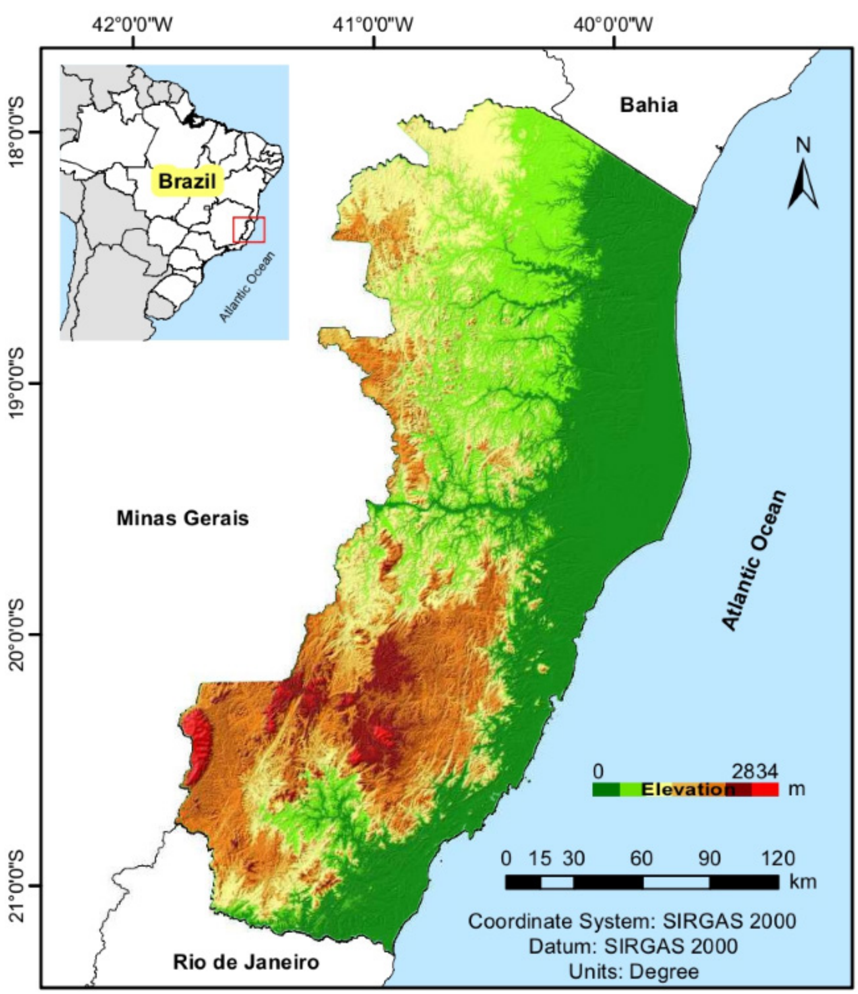

Figure 1. Geographical location of the study area with digital elevation model.

in the state southwest region; Am, humid or subhumid tropical climate, found in the state northeastern; and Aw, tropical climate with dry winter, found in the state Western region.

The Atlantic Forest is among the most biodiverse and threatened regions on the planet (Myers et al. 2000). Since colonization, the Brazilian Atlantic Forest has been suffering from constant processes of forest fragmentation arising from the different cycles of land occupation (Dean 1995), where the alteration of the primitive landscape by humans is thousands of times greater than the dynamics of natural disturbance of the ecosystem (Tabarelli & Gascon 2005). Consequently, this fragmentation process reduced the original area of the Atlantic Forest biome to the current 12.4%. The entire Espírito Santo State is covered by this biome, with about 10.5% of the original remaining (SOS Mata Atlântica 2019). Therefore, the remnants inspire greater attention concerning conservation policies, especially those related to implantation and management of protected areas.

The following sections describe data set preparation and used methodology. First of all, kernel density and fire maps are explained. Then, CART theory and the requirements for its implementation are illustrated. Finally, prediction map is described, paying particular attention on how predictors are implemented. All methodological approach is described and summarized in Figure 2. Flowchart shows the main procedures involved in the process of determining fire density and making fire predictions.

Dataset

Most researches related to fire occurrence are based on fire records and requires data covering several fire stations. The occurrence of fire, which is maintained by firefighting agencies, is the most common source of such data. Although occasionally, when these records are not available, satellite images can be used (Dlamini 2010, Maingi & Henry 2007, Prasad et al. 2008).

MCD45A1 product maps fire-affected areas, is monthly distributed and belongs to MODIS 5 collection (Roy et al. 2008). MCD45A1 subsets, in shapefile format and ranging from 2000 to 2015, have been downloaded via Server File Transfer Protocol (FTP) from the website http://modis-fire.umd.edu/ by software SmartFTP (1 of Figure 2). Files are available in sinusoidal projection Lat-Long with geographic extension in subcontinental windows. Study area is bounded by the 6 window, covering Central South America with latitude S between 10° and 35° and longitude W between 34° and 79°. Monthly maps of burned area were then projected to Universal Transverse Mercador (UTM), datum SIRGAS 2000 and converted to raster format (2 of the Figure 2) with spatial resolution of 250m.

The availability of fire records has restricted the search for fire occurrences in the past. The few studies as Santos el al. (2006) reports that the main causes of fires in the national territory protected areas are associated with human action, mainly in the category of incendiary and burning for cleaning, with the most frequent occurrences from July to October. In Espírito Santo State, information on the profile of forest fires in protected areas can be found in Tebaldi et al. (2013). Since the advance of Earth observation from space, remote sensing has become a valuable tool for the scientific community and natural resource managers. The mapping of the burned area by global satellite systems has become essential for development of environmental management policies. For the considered study period, the historical records of fire detected by remote sensing in the Espírito Santo State registered more than 20,000 hectares of burnt area. This has caused

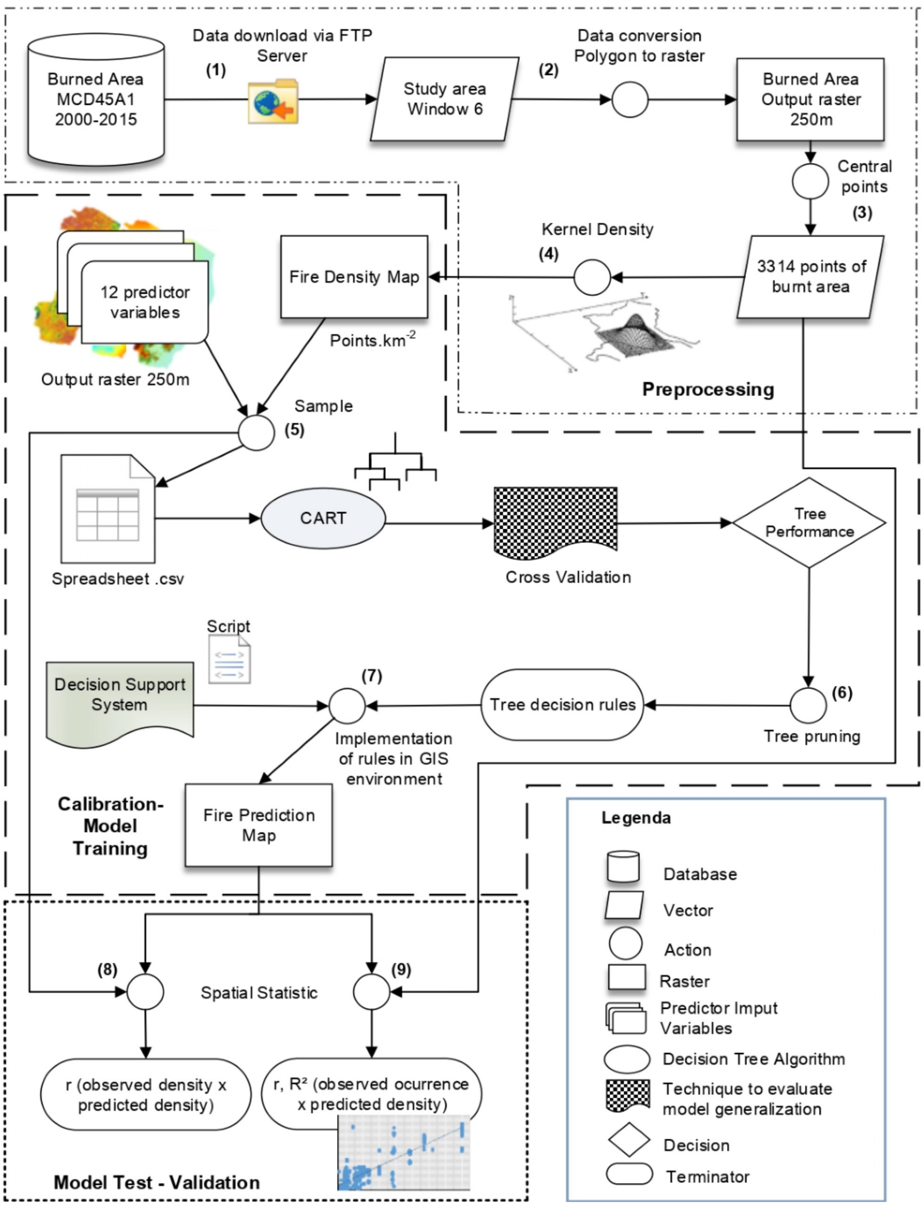

Figure 2. Overview of the main procedures involved in the process of determining fire density and making fire predictions.

the reduction of forest fragments and loss of biodiversity, mainly in the state conservation units (Tebaldi et al. 2013).

Kernel Density

Center points (3 of the Figure 2) of each pixel from the raster image were considered to spatialize the burned area data, totaling 3314 points. Interpolation techniques can be used to convert punctual to continuous data, in order to predict attribute values for not sampled locations (Burrough & McDonnell 1998). In case of finite point observations, results from kernel density estimates are suitable (Bowman & Azzalini 1997). This approach was originally developed as an alternative method to obtain a smooth probability density function, univariate or multivariate, from a sample of observations (Bailey & Gatrell 1995, Levine 2002). As estimated intensity of punctual observations (coordinates given in *x* and *y*) is very similar to bivariate probability density, kernel approach can be adapted for this purpose (Bailey & Gatrell 1995).

Kernel density estimation (4 of the Figure 2) is a non-parametric statistical method for estimating probability densities. One kernel (that is, bivariate normal probability density) is placed on each observation point and the crossing intensity of an overlapping grid is estimated (Seaman & Powell 1996). The method is similar to "mobile window" concept, in which a specific sized window is moved over the observation points (Gatrell et al. 1996). Mathematically (Parzen 1962, Rosenblatt 1956), for a location with vector coordinates x_j, density $\hat{f}(x)$ can be expressed by the following equation (Eq. 1):

$$\hat{f}(x) = \frac{1}{nh^2}\sum_{i=1}^{n} K\left\{\frac{(x_j - X_i)}{h}\right\} \tag{1}$$

Where: n is the number of observations; x_i is the vector of fire coordinates; K is the kernel function; h is the search radius or bandwidth.

Various functions of K interpolation, differ in the way that they assign weights to points within a search radius, which can be any probability density function (Gaussian, triangular, quartic, uniform or negative exponential) that meets the following equality (Eq. 2):

$$\int_{-\infty}^{+\infty} K(h)\,dh = 1 \tag{2}$$

Quartic kernel function (Silverman 1986) was considered for calculation of function K in sofware ArcGis/ArcInfo 10.4 (Eq. 3):

$$k(h) = \frac{3}{\pi}\left(1 - h^2\right) \tag{3}$$

Quartic function ponders with greater weight distant points with gradual decrease. Search radius or bandwidth expresses kernel size and controls smoothing on the generated surface. Fixed and adaptive methods can be applied to estimate kernel density. In the fixed method, the search radius is defined in units of distance and is constant throughout the interest area. In adaptive method, the search radius is defined by the minimum number of individual observations found in the kernel and depends on concentration of punctual observations. This means that, in areas of low concentration, search radius has higher values than in areas of high concentration (Worton 1989).

An important question that is hard to define when implementing kernel density interpolation, in both fixed and adaptive methods, is the choice of kernel smoothing parameter. A smaller search radius allows close observations dominate density estimate, while larger search radius favor distant observations (Worton 1989, Seaman & Powell 1996). In the literature, some different methods have been proposed to set smoothing parameters in order to evaluate occurrence and pattern of fires (Amatulli et al. 2007, de la Riva et al. 2004, Koutsias et al. 2004, Liu et al. 2010). However, the choice of an arbitrary value for the smoothing parameter is not recommended and should be done in a more rigorous way, so that the model is not penalized. Therefore, search radius h was calculated using a space variant of Silverman (1986) that is robust to outliers (i.e. points that are far apart from other points) and it is implemented in ArcGis/ArcInfo 10.4 software. Thus, outlined the kernel density configuration of the model, the surface map of fire occurrence, with grid resolution of 250 cells, was generated and used in regression tree analysis as the response variable. Therefore, this spatial resolution allows to eliminate the error concerning the fire starting locations.

Predictive variables

According to literature and taking into consideration the significance of each variable to explain forest fires occurrence within study area, a total of 12 variables were considered

(Table I), including aspects topographical (Supplementary Material - Figure S1), climatic (Figure S2), socioeconomic (Figure S3) and vegetation (Figure S4). Variables from various databases were then processed into raster image, with spatial resolution of 250m, in accordance with procedures described in Supplementary Material. Geographical data were configured in accordance with geocentric reference system (SIRGAS 2000) and integrated into ArcGis/ArcInfo 10.4 GIS environment.

Training model - calibration

With fire density response variable along with the entire database of predictors, variable sample tool (5 in Figure 2) of ArcGis/ArcInfo 10.4 was used for systematic sampling and data input on CART decision tree algorithm. Systematic sampling ensures a large amount of data for tree algorithm training/calibration and test/validation.

CART decision tree algorithm (Breiman et al. 1984) is a recursive binary partitioning procedure

Table I. Predictive variables considered in the decision tree analysis.

Variable		Variable file name	Grasp	Unit
Topographic	Altitude	DEM	0 – 2834	m
	Slope	SLO	0 – 85.40	°
	Composite topographical index	CTI	0.75 – 16.07	adimensional
Environmental	Annual average rainfall	PREC	885.94 – 1,817.98	mm
	Annual average temperature	TEMP	6.54 – 25.79	°C
	Solar radiation	SOLAR_RAD	0.23 – 1.30	MJ cm^{-2} hr^{-1}
	Annual average water deficit	DEF_HID	0 – 603	mm
Socioeconomic	Demographic density	DEMOG_ DENS	0 – 5,2397.7	hab km^{-2}
	Income	INCOME	0 – 2,7017.04	R$ $month^{-1}$
	Proximity to roads	PROX_ROADS	0 – 4,854.12	m
Vegetation	Vegetation continuous fields	VCF	0 – 86	%
	Land use and cover	1-Agriculture 2-Urban areas 3-Water course 4-Natural forest 5-Mangrove 6- Pasture 7-Forestry 8-Exposed soil 9- Flooded areas 10-Restinga		

capable of processing continuous and nominal attributes as predictors and targets. Data are treated in raw form; no binning is necessary nor recommended. Starting at root node, data are divided until terminal nodes, without a stopping rule (Wu & Kumar 2009). Algorithm starts by analyzing all input variables and determines which binary division of a single predictor best reduces deviation in the response variable. This process is repeated for each partition of data resulting from first division, continuing until homogeneous terminals nodes are achieved in the hierarchical tree. The technique typically causes model overfitting, creating a tree that explains substantially all deviation in the original data. Then the tree has to be pruned back by methods of cost complexity (for more details see Esposito et al. 1997). Pruning method involves cross-validation (Venables & Ripley 2002), which consists of equally dividing the original data set; which will be used to generate test trees that will be validated against the last set. Estimates of mean squared error from cross validation help to select the most convenient tree size, considering a trade-off between reduction in expected error and convenience of generating a reasonable amount of decision rules.

CART engine includes automatic handling missing data, the construction of dynamic resources (Wu & Kumar 2009), which is robust to outliers and do not require a priori variables selection. In addition, it can model the relationships among variables, despite significant spatial autocorrelation (Cablk et al. 2002), as well as the relative importance of each variable used in the model (Steinberg & Colla 1997).

Training database was used to implement CART regression tree algorithm using demo version Salford Predictive Modeler (SPM) 8.0 software. In this study, a built-in validation/calibration procedure of tree's performance was conducted by a 10-fold cross-validation, which was able to produce validation errors for each generated tree. First, a large tree was generated aiming division of least squares; successively, the greater tree was pruned (6 in Figure 2) to get a good level of cross-validation error, allowing selection of a smaller tree. Amatulli et al. (2006) mentions that one large tree can produce a very detailed regression process, creating decision rules for small units of fire risk. As a result, algorithm would then be too complex, reducing its interpretive nature. In addition, small units increase meaningless fire planning segmentation, thereby losing its operational efficiency. In general, a satisfactory level of cross-validation error and unit size must be identified, considering interpretability of decision rules.

Decision rules, based on threshold values of a specific predictor variable, were implemented in python, in order to read Excel file database and determine output tree values. Output values were then imported into GIS environment (7 in Figure 2), to map predicted fire density of each grid cell, allowing creation of the final map of fire prediction. This combination provides an important tool to spatially locate preventive actions that must be taken, within a framework of fire management.

Model-calibration test

A validation of the resulting map was obtained analyzing the correlation between predicted risk values and observed fire values, through r coefficient (8 in Figure 2).

Subsequently, in order to check model predictive capacity and the influence of fire point position in the study area, a second correlation analysis was performed, in which points of burns in each fire management unit, expressed in observed density, were graphically represented as a function of predicted density values of fire risk map (9 in Figure 2). Then, obtained regression equation with its slope, intercept and Pearson coefficient was calculated. In addition, through CART analysis the relative importance of each variable in regression process was set, revealing their capability for predicting fire risk.

RESULTS

Fire density map

Kernel map of fire density (Figure 3) highlighted four main areas of greatest burning. One is in State's northeast region, with peaks ranging from 0.2 to 0.6 points km^{-2}. Another is in Rio Doce region, with more outstanding peaks, ranging from 0.3 to 0.9 points km^{-2}; and, two other areas in the south coast and Caparaó regions, with peaks ranging from 0.2 to 0.3 points km^{-2}. In the other regions from study area, the occurrence of fire slightly decreases to values from 0 to 0.2 points km^{-2}.

Decision tree rules

The large dataset and number of predictor variables created a very complex tree, with 5.137 terminal nodes and a cross-validation error of 0.01. Tree pruning was successively performed to obtain a simple tree with 38 terminal nodes (Figure 4) and a cross-validation error with acceptable values (0.3). Tree decision rules identified several unique thresholds for each variable and fire management unit, allowing to predict 38 average densities, ranging from 0.004 to 0.748 points km^{-2}, in order to smooth maximum values (0.9 points km $^{-2}$) of the fire density map.

Management units range from large (199.101 grid cells) to small (179 grid cells). Lower risk zones represent areas of greater size within state (Figure 5), while zones of greatest risk in the prediction map correspond to coastal and northeast state regions, according to the fire density map.

Spatial validation of the map indicates an acceptable correlation (0.82), confirming that the selected tree was able to provide a reliable fire prediction map, following the trend available on the fire density map. In addition, CART algorithm generated an acceptable regression model, with an adjusted coefficient of determination of 0.88 and a correlation of 0.94 between predicted density on fire risk map and observed density at points of burning area. In general, the lowest values of density in the risk map present better fit in the model (Figure 6).

Table II shows each variable score in the decision tree process. The variable with highest predictive capacity was population density (100.0), followed by variable rainfall (78.4) and land use and cover (75.07). Income (36.05) and altitude (33.51) also had good predictive capacity in the model, while continuous field (1.07) and solar radiation (0.01) were not significant in the model development. In general, socioeconomic, environmental and vegetation factors are more important on regional scale fire prediction, once they present higher punctuation, which is

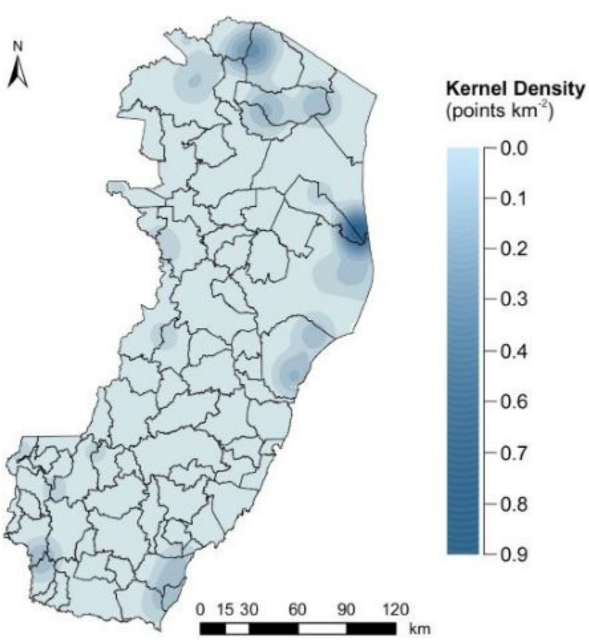

Figure 3. Kernel density map of fire.

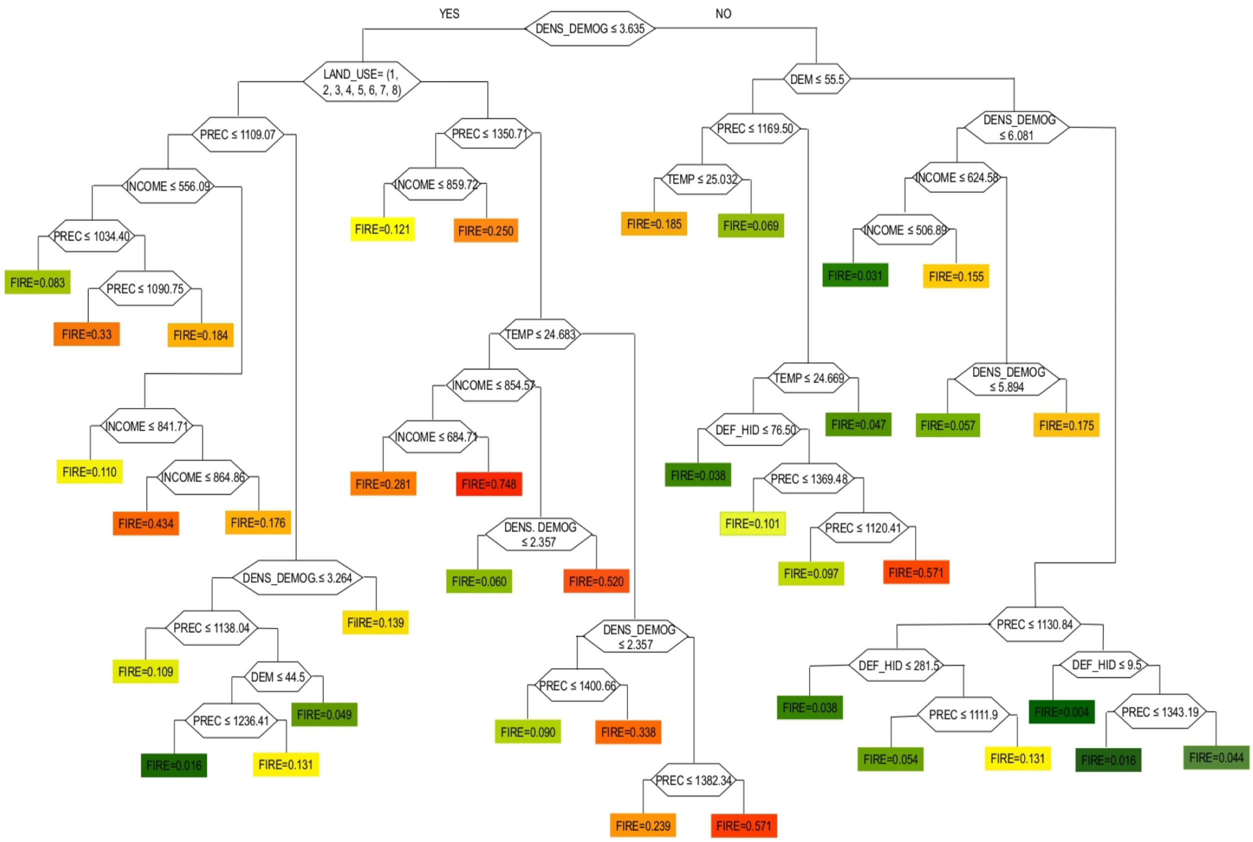

Figure 4. **Regression rules described in binary tree form. Predicting variables codes and their classes or ranges are listed in Table I.**

confirmed by their abundant presence in the tree structure, for all fire density scales (Figure 4). Topographic variables, on the contrary, are less important in the predictive model. Among these, altitude is the most relevant variable. Finally, proximity to roads variable shows less influence on the regression process when compared to above-mentioned variables.

DISCUSSION

Fire Density Map

Information technology has become important for monitoring burning areas as well as predicting occurrence and behavior of forest fires. Improvements in support systems and data quality resulted in more efficient decision making for fire response and forest management. Development of spatial statistical models led to notable improvements in fire predictive capacity, through integrating fire risk classification system with information application and space technology. Methods such as kernel density analysis provide an instrument for forest managers to develop maps of fire occurrence in situations of spatial and temporal variability.

Kernel density technique is often used for different ecological applications, such as area analysis of life (Millspaugh et al. 2006) and landscape fragmentation studies (Cai et al. 2013). Choosing an appropriate smoothing parameter (that is, bandwidth) is the most important step in obtaining a kernel density estimator (Worton 1989), but there is no agreement on how to address this problem (Downs & Horner 2007, Fieberg 2007, Gitzen et al. 2003, Horne & Garton 2006).

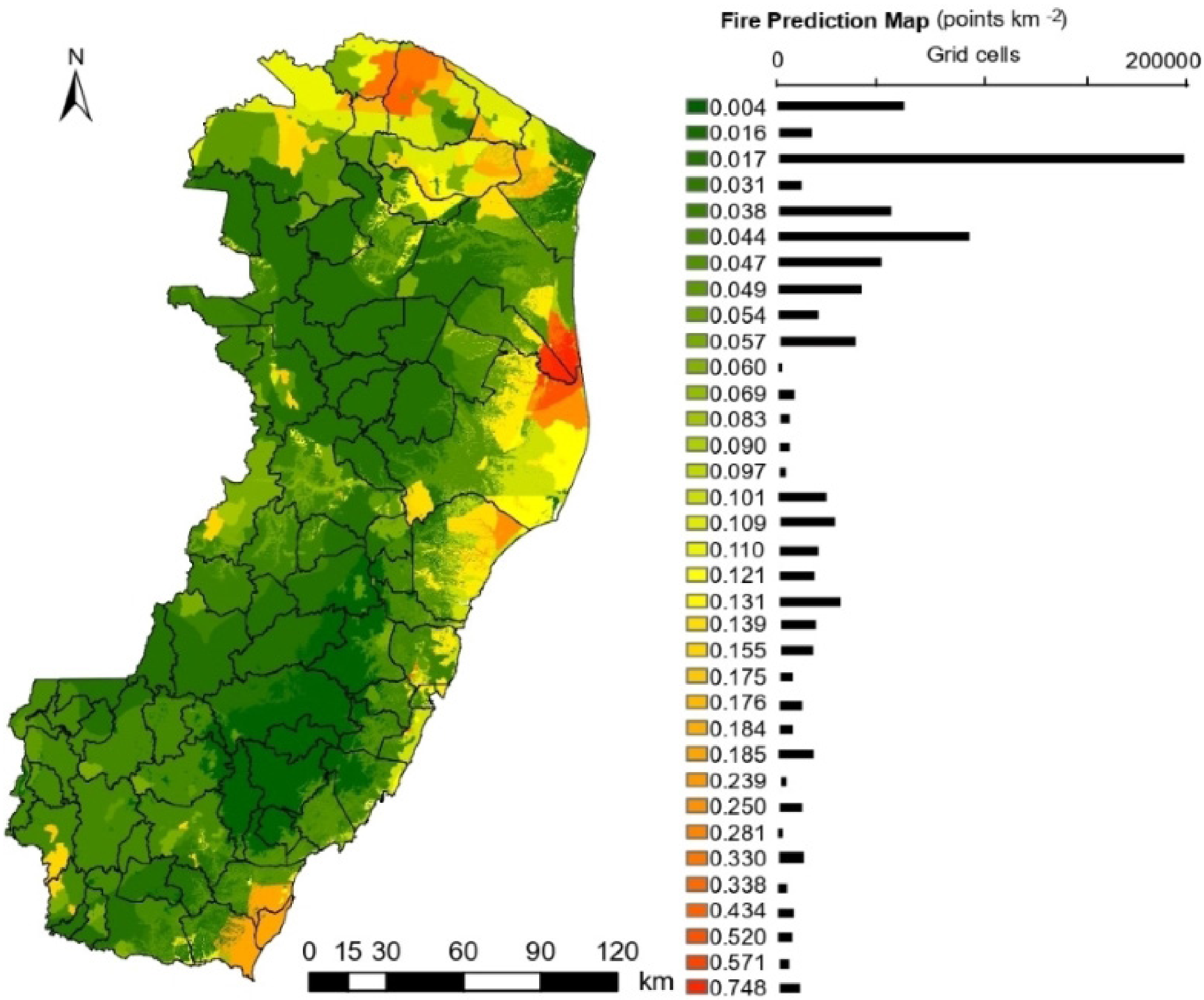

Figure 5. Fire prediction map obtained by application of decision rules.

Smoothing parameter (h) determines the kernel propagation centered on each observation. If h value is small, individual kernels will be narrow and kernel density estimate at a given point will be based on only few observations. This may not allow variation between samples and may produce a poorly smoothed map (high values). On the other hand, if h value is large, individual kernels will be large as well, which can hide fine details and result in a very smoothed map (low values). Therefore, the smoothing approach to use depends on the set of observations along with ecological considerations specific to each study/purpose. The spatial pattern of fire distribution in the study area has a grouped behavior. Its occurrence in specific areas will depend on several factors related to the legal protection of natural resources, properties and forest management.

Decision tree rules

The decision tree model generated by CART algorithm, presents advantages over some techniques used in the context of forest fire prediction. It is a flexible model and allows the user to define a tree architecture with a reduced number of terminal nodes, with good

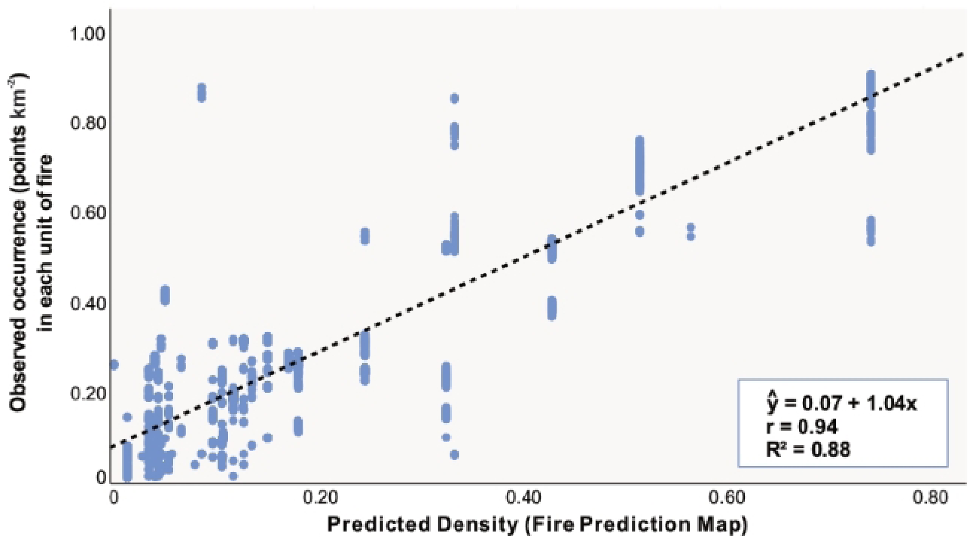

Figure 6. Regression equation obtained between predicted density and observed occurrence of fires in each fire management unit.

predictability for forest fire risk areas. Although, other approaches, such as random forest, work with the same principle, CART algorithm stands out for the practicality of creating a single tree that is easy to interpret. Thus, by knowing the variables present in the tree structure along with their respective values, that determine the risk areas for forest fires, an effective planning can be elaborated for the different regions in the study area.

More precise conclusions can be reached based on the structure of the tree algorithm. High-density fire zones are associated with either a high or low population density. Areas of restinga vegetation and flooded areas are generally associated with high fire density (0.748). Thus, some management strategies can be performed in such areas as restricting access to risky locations and managing fuel to prevent onset and spread of fire. An important application of the algorithm is the possibility of data self-feeding, in order to automatically develop fire prediction. In addition, there is the possibility of creating scenarios based on simulated changes in the data, aiming to observe new spatial dispositions of fire management units and their risk value.

An important role is given to the climatic parameter represented by rainfall, temperature and water deficit variables. This parameter is an important ecological indicator, not only in species composition definition, but also in their distribution. Such variables are capable of discriminating ecological conditions and fire susceptibilities not detected by land cover data. Species susceptibility to forest fires is not only related to tree species flammability and plantation structure, but also to state of water stress, which is directly influenced by average meteorological conditions (Aguado et al. 2003, Chuvieco & Martin 1994). Predictions of fire regimes assume a strong link between climate and fire, but generally with less emphasis on effects of local factors such as human activity (Liu et al. 2010, Wotton et al. 2010).

The importance of social factors can also be pointed out by the income variable, since forest fires are associated with lower income sites. In the USA, forest fires that start in poor communities are less likely to extinguish out rapidly due to resource lacking (Mercer & Prestemon 2005). Ecological damage resulting from forest fires can harm the natural resource base from which communities originate their economic activities (Butry et al. 2001). These

Table II. Punctuation of each variable in the decision tree process

Variable	Punctuation (%)
Demographic density	100
Annual average rainfall	78.4
Land use and cover	75.07
Income	36.05
Altitude	33.51
Composite topographical index	24.06
Annual average water deficit	22.89
Annual average temperature	22.13
Proximity to roads	9.92
Declivity	9.14
Continuous field of vegetation	1.04
Solar radiation	0.01

findings suggest that social conditions can be key determinants of social vulnerability and fire risks. Understanding how these vulnerability components vary can help managers develop appropriate mitigation and protection strategies for specifics sites and populations.

Some observations can be extracted from the model. According to Amatulli et al. (2006), a minimum or maximum fire management unit size can be defined based on prevention actions and fire planning guidelines. Predefined fire management unit sizes should be used in tree growing phase in order to reinforce regression rules and group homogeneous pixels into areas larger or smaller than the defined dimension. This process would also allow an evident reduction in the tree size, improving understanding of the whole regression process.

Prediction Fire Map

Spatial location of fire management units and study area knowledge in terms of ecology and socioeconomic factors can be important indicators of fire causes. In the micro region of Rio Doce (linhares) and Northeast (São Mateus and Linhares municipalities), close to the coast of the state (Figure 5), where a greater prediction of forest fires was observed, the risk fire is related to two main causes. First is due to the use of fire as a management resource for farm cultivation. These areas are characterized by a high level of fragmentation (Juvanhol et al. 2021). These results are consistent, according to work done in tropical forests (Cochrane & Laurance 2002, Holdsworth & Uhl 1997), where edges between wildland - urban interface are considered to be the most vulnerable to forest fires, as it also occurs in temperate forests (Ganteaume et al. 2013, Maselli et al. 2003, Yang et al. 2007). Landscape fragmentation level strongly widens boundary borders, increasing the likelihood that human activities will affect ecological processes in natural areas; in this specific case, increasing vulnerability to fire (Juvanhol et al. 2021; Leone & Lovreglio 2003). Second cause is more related to vegetation, being the transition range between forest and restinga. Predominance of herbaceous plants established in sandy soil with a high concentration of organic matter, facilitates fire ignition (Juvanhol et al. 2021).

Tropical rainforests of the Atlantic Forest, despite their location in one of the wettest areas of Brazil, where average annual rainfall is over 1500mm, sporadically suffer from fires (Oliveira & Passacantili 2010). According to Hammond et al. (2007), humid forests often suffer from fire in Guianas due to anthropic impacts. Carcaillet et al. (2002) report that in the last 2000 years, in the Amazon, fires are also present in the paleoenvironmental record, also associated with anthropic influence. More recent studies in temperate forests have found similar results regarding transformation of large forests into open or low vegetation areas throughout altered fire regimes, in face of climate and land use change (Paritsis et al. 2015, Tepley et al. 2016).

Understanding confluence of social and biophysical vulnerability is especially relevant to forest fires. Frequency, severity, and pattern of forest fires are significantly related to human activities, including land use, population establishment patterns, and vegetation management (Hawbaker et al. 2013; Syphard et al. 2007, 2013). For example, the occurrence of forest fires is positively associated with population and housing density (Syphard et al. 2007, Hawbaker et al. 2013), once people cause most of fire ignition, land use influences vegetation patterns and, therefore, fire behavior (Prestemon et al. 2013).

While some policies, such as those supporting research on sustainable forms of agriculture, have been introduced in recent years, forest areas continue to receive less favorable treatment of rural taxation schemes. Above all, political and institutional structure still favors extensive cultivation practices and land conversion (Wigtil et al. 2016)

We must pay more attention to development of appropriate conservation actions. The specific need to strengthen efforts in areas where land protection is already established stands out. In addition, areas that provide vital environmental services to local communities, where impacts and threats are particularly concentrated, must also be protected.

In order to better understand the effects of management actions on fire occurrence and behavior, field studies should be conducted at various sites and at various scales. Some federal policies indirectly support individuals in reducing their vulnerability to fire. The 2003 Healthy Forests Restoration Act in EUA provides opportunities to develop community-based forest fire protection plans to enhance their ability to adapt to forest fires (Grayzeck-Souter et al. 2009, Jakes et al. 2011, Williams et al. 2012). These plans have been shown to improve community resilience (Jakes & Sturtevant 2013). Development of community forest fire protection plans in vulnerable communities can help to reduce their susceptibility to fire impacts, once they depend on communities that have access to suitable resources (Jakes et al. 2011).

Study implications

The results presented by the regression tree model are very interesting when compared to other works carried out at the local and national level (Arpaci et al. 2014, Bem et al. 2018, dos Santos et al. 2020, Liang et al. 2019, Oliveira et al. 2012, Zhang et al. 2016). Particularly, within European major fire projects, the global precision achieved through logistic regression and neural networks was 60% and 69%, respectively (Bart 1998, Chuvieco 1999).

In the context of the Decision Support System (DSS), the relevance of the study stands out for allowing the forest manager to choose the priority conservation regions according to the local environmental resilience and valuation. In this sense, a differentiated operational plan can be carried out in each fire zone, taking into account the survey of operating costs on local and regional scales.

Other approaches in the context of forest fire risk modeling were employed with a good performance of the prediction maps (Martín et al. 2019, Zhang et al. 2019). Even though, these models allow calculating the variables importance scores to determine the global map of the area. The proposal presented here is more complete in the context of DSS, once the model also allows to understand how the variables are related and locally influence the fire management areas. Thus, the forest manager can get insights about the causes of fire occurrences and consider this information during decision making.

Additionally, the importance of an attribute in CART model is applied to the total improvement of all nodes in which the attribute appears as a divisor (weighted by the fraction of the training data in each node division). The substitute divisor is also included in the importance calculations, which means that even a variable that does not divide a node can receive a score of great importance (Wu & Kumar 2009). Substitutes are an important innovation in machine learning and play a key role in predicting and interpreting the CART tree. When a divisor has no close substitutes, it means that the information content of that variable is unique and cannot be easily replaced by any other variable.

Further research is needed to extend model implementation and validation, using other study sites, and a more extensive set of predictor variables. The same methodology can even be applied in the short term fire risk assessment framework.

Although, this advanced data analysis was focused on fire risk assessment, it can be extended to other fields of science. In particular, to natural disasters risk, where several factors are generally involved. Their nature and behavior is often not well known and multidisciplinary interaction is needed in order to emphasize the complex mechanism of their possible relationships.

CONCLUSIONS

This study demonstrated that two non-parametric techniques, combined with GIS, can provide a significant model for predicting units of fire risk. The resulting decision tree model shows good performance between its dimension and cross-validation error. In general, socioeconomic, environmental and vegetation factors have a greater importance in the proposed prediction model.

Proposed Decision Support System (DSS) provides a sound basis in the overall context of risk analysis. More emphasis should be placed on applying non-parametric techniques, once they can correctly address a wide range of environmental issues in order to explore data distribution and intrinsically variable relationships.

Acknowledgments

The authors would like to thank Thiago Tuler for his fundamental help in implementing the decision tree rules. Thanks also to Espírito Santo Research and Innovation Support Fundation (FAPES) – Notice 001/2014 – Grant Term Number: 179/2014. To United States Geological Survey (USGS), Instituto Nacional de Meteorologia – INMET (National Institute of Meteorology), Instituto de Meio Ambiente e Recursos Hídricos – IEMA (Institute of Environment and Water Resources), Sistema Integrado de Bases Geoespaciais do Estado do Espírito Santo – GEOBASES (Geospatial Integrated System of Espirito Santo State), Instituto Brasileiro de Geografia e Estatística – IBGE (Brazilian Institute of Geography and Statistics), for providing part of the database. To the researchers that compose the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) research group Geotechnology Applied to Global Environment (GAGEN) and the Pos-Graduate Program in Forest Sciences of the Federal University of Espirito Santo.

REFERENCES

AGUADO I, CHUVIECO E, MARTIN P & SALAS J. 2003. Assessment of forest fire danger conditions in southern Spain from NOAA images and meteorological indices. Int J Remote Sens 24(8): 1653-1668.

AMATULLI G, PERÉZ-CABELLO F & DE LA RIVA J. 2007. Mapping lightning/human-caused wildfires occurrence under ignition point location uncertainty. Ecol Model 200(3-4): 321-333.

AMATULLI G, RODRIGUES MJ, TROMBETTI M & LOVREGLIO R. 2006. Assessing long-term fire risk at local scale by means of decision tree technique. J Geophys Res 111(4): 1-15.

ANDREAE MO & MERLET. 2001. Emission of trace gases and aerosols from biomass burning. Global Biogeochem Cy 15(4): 955-966.

ARGAÑARAZ JP, PIZARRO GG, ZAK M, LANDI MA & BELLIS LM. 2015. Human and biophysical drivers of fires in Semiarid Chaco mountains of Central Argentina. Sci Total Environ 520: 1-2.

ARPACI A, MALOWERSCHNIG B, SASS O & VACIK H. 2014. Using multi variate data mining techniques for estimating fire susceptibility of Tyrolean forests. Appl Geogr 53: 258-270.

BAILEY TC & GATRELL AC. 1995. Interactive spatial data analysis. Harlow: Longman, 432 p.

BART LC. 1998. Evaluación de la estimación de grandes incendios forestales en la cuenca mediterránea Europea por redes neuronales y regresión logística. Série Geográfica 7: 73-85.

BEM PP, JÚNIOR OAC, MATRICARDI EAT, GUIMARÃES RF & GOMES RAT. 2018. Predicting wildfire vulnerability using logistic regression and artificial neural networks: a vase study in Brazil's Federal District. Int J Wildland Fire 28(1): 35-45.

BONAZOUNTAS M, KALLIDROMITOU D, KASSOMENOS PA & PASSAS N. 2005. Forest fire risk analysis. Hum Ecol Risk Assess 11(3): 617-626.

BOWMAN AW & AZZALINI A. 1997. Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrattions. New York: Publications, Oxford Science, 193 p.

BREIMAN L, FRIEDMAN JH, OLSHEN RA & STONE CJ. 1984. Classification and Regression Trees. Chapman & Hall, 368 p.

BURROUGH PA & MCDONNELL RA. 1998. Principles of Geographical Information Systems. Oxford University Press, 356 p.

BUTRY DT, MERCER ED, PRESTEMON JP, PYE JM & HOLMES TP. 2001. What is the price of catastrophic wildfire? J Forest 99(11): 9-17.

CABLK M, WHITE D & KIESTER AR. 2002. Assessment of spatial autocorrelation in empirical models in ecology. In: Scott JM et al. (Eds), Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, DC. p. 429–440.

CAETANO M, FREIRE S & CARRÃO H. 2004. Fire risk mapping by integration of dynamic and structural variables. Remote Sensing in Transition 1: 319-326.

CAI X, WU Z & CHENG J. 2013. Using kernel density estimation to assess the spatial patterns of road density and its impact on landscape fragmentation. Int J Geogr Inf Sci 27(3): 222-230.

CARCAILLET C ET AL. 2002. Holocene biomass burning and global dynamics of the carbon cycle. Chemosphere 49(8): 845-863.

CATRY FX, REGO FC, BAÇÃO FL & MOREIRA F. 2009. Modeling and mapping wildfire ignition risk in Portugal. Int J Wildland Fire 18(8): 921-931.

CHUVIECO E. 1999. Remote Sensing of Large Wildfires in the European Mediterranean Basin. Berlim: Springer, 212 p.

CHUVIECO E & CONGALTON RG. 1989. Application of Remote Sensing and Geographic Information Systems to Forest Fire Hazard Mapping. Remote Sens Environ 29(2): 147-159.

CHUVIECO E & MARTIN MP. 1994. Global fire mapping and fire danger estimation using AVHRR images. Photogramm Eng Rem S 60(5): 563-570.

COCHRANE MA & LAURANCE WF. 2002. Fire as a large-scale edge effect in Amazonian forests. J Trop Ecol 18(3): 311-325.

COVINGTON WW & MOORE MM. 1994. Southwestern Ponderosa Forest Structure: Changes since Euro-American Settlement. J Forest 92(1): 39-47.

DE LA RIVA J, PÉREZ-CABELLO F, LANA-RENAULT N & KOUTSIAS N. 2004. Mapping wildfire occurrence at regional scale. Remote Sens Environ 92(2): 288-294.

DEAN W. 1995. With broadax and firebrand: the destruction of the Brazilian Atlantic Forest. San Francisco: University of California Press, 504 p.

DLAMINI WM. 2010. A Bayesian belief network analysis of factors influencing wildfire occurrence in Swaziland. Environ Modell Softw 25(2): 199-208.

DOS SANTOS EE, SENA NC, BALESTRIN D, FERNANDES FILHO EI, COSTA LM & ZEFERINO LB. 2020. Prediction on Burned Areas Using the Random Forest Classifier in the Minas Gerais State. Floresta e Ambient 27(3): e20180115.

DOWNS JA & HORNER MW. 2007. Effects of Point Pattern Shape on Home-Range Estimates. J Wildlife Manage 72(8): 1813-1818.

ESPOSITO F, MALERBA D & SEMERARO G. 1997. A Comparative analysis of methods for pruning decision trees. IEEE Trans Pattern Anal Mach Intell 19(5): 476-491.

EUGENIO FC, DOS SANTOS AR, FIEDLER NC, RIBEIRO G, DA SILVA AG, DOS SANTOS ÁB, PANETO GG & SCHETTINO VR. 2016. Applying GIS to develop a model for forest fire risk: A case study in Espírito Santo, Brazil. J Environ Manage 173: 65-71.

FIEBERG J. 2007. Kernel density estimators of home range: smoothing and the autocorrelation red herring. Ecology 88(4): 1059-1066.

FONSECA MG, ARAGÃO LEOC, LIMA A, SHIMABUKURO YE, ARAI E & ANDERSON LO. 2016. Modelling fire probability in the Brazilian Amazon using the maximum entropy method. Int J Wildland Fire 25: 955-969.

GANTEAUME A, CAMIA A, JAPPIOT M, SAN-MIGUEL-AYANZ J, LONG-FOURNEL M & LAMPIN C. 2013. A Review of the Main Driving Factors of Forest Fire Ignition Over Europe. Environ Manage 51(3): 651-662.

GATRELL AC, BAILEY TC, DIGGLE PJ & ROWLINGSON BS. 1996. Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. T I Brit Geogr 21(1): 256-274.

GITZEN RA, MILLSPAUGH JJ & KERNOHAN BJ. 2003. Bandwidth Selection for Fixed-Kernel Analysis of Animal Utilization Distributions. J Wildlife Manage 70(5): 1334-1344.

GONZÁLEZ JR, PALAHÍ M, TRASOBARES A & PUKKALA T. 2006. A fire probability model for forest stands in Catalonia (north-east Spain). Ann For Sci 63(2): 169-176.

GRAYZECK-SOUTER SA, NELSON KC, BRUMMEL RF, JAKES P & WILLIAMS DR. 2009. Interpreting federal policy at the local level: the wildland – urban interface concept in wildfire protection planning in the eastern United States. Int J Wildland Fire 18(3): 278-289.

HAMMOND DS, STEEGE HT & VAN DER BORG K. 2007. Upland Soil Charcoal in the Wet Tropical Forests of Central Guyana. Biotropica 39(2): 153-160.

HARDY CC, OTTMAR RD, PETERSON JL, CORE JE & SEAMON P. 2001. Smoke management guide for prescribed and wildland fire: 2001 edition. Boise, ID: National Wildfire Coodination Group, 226 p.

HAWBAKER TJ, RADELOFF VC, STEWART SI, HAMMER RB, KEULER NS & CLAYTON MK. 2013. Human and biophysical influences on fire occurrence in the United States. Ecol Appl 23(3): 565-582.

HOLDSWORTH AR & UHL C. 1997. Fire in amazonian selectively logged rain forest and the potential for fire reduction. Ecol Appl 7(2): 713-725.

HORNE JS & GARTON EO. 2006. Likelihood Cross-Validation Versus Least Squares Cross- Validation for Choosing the Smoothing Parameter in Kernel Home-Range Analysis. J Wildlife Manage 70(3): 641-648.

JAKES PJ ET AL. 2011. Community wildfire protection planning: is the Healthy Forests Restoration Act's vagueness genius? Int J Wildland Fire 20(3): 350-363.

JAKES PJ & STURTEVANT V. 2013. Trial by fire: Community wildfire protection plans put to the test. Int J Wildland Fire 22(8): 1134-1143.

JUVANHOL RS, FIEDLER NC, SANTOS AR, SILVA GF, OMENA MS, EUGENIO FC, PINHEIRO CJG & FERRAZ FILHO AC. 2021. Gis and fuzzy logic applied to modelling forest fire risk. An Acad Bras Cienc 93: e20190726. https://doi.org/10.1590/0001-3765202120190726.

KALABOKIDIS KD, KOUTSIAS N, KONSTANTINIDIS P & VASILAKOS C. 2007. Multivariate analysis of landscape wildfire dynamics in a Mediterranean ecosystem of Greece. Area 39(3): 392-402.

KOUTSIAS N, KALABOKIDI, KD & ALLGÖWER B. 2004. Fire occurrence patterns at the landscape level: beyond positional accuracy of ignition points with kernel density methods. Nat Resour Model 17(4): 359-375.

LENTILE LB, HOLDEN ZA, SMITH AM, FALKOWSKI MJ, HUDAK AT, MORGAN P & BENSON NC. 2006. Remote sensing techniques to assess active fire characteristics and post-fire effects. Int J Wildland Fire 15(3): 319-345.

LEONE V & LOVREGLIO R. 2003. Human fire causes: A challenge for modelling. In: Chuvieco E et al. (Eds), 4th Intern. Workshop on Remote Sensing and GIS Applications to Forest Fire Management: Innovative Concepts and Methods in Fire Danger Estimation, p. 89-94.

LEVINE N. 2002. Crime Stat II - A Spatial Statistics Program for the Analysis of Crime Incident Locations. The National Institute of Justice, Washington, DC.

LIANG H, ZHANG M & WANG H. 2019. A neural network model for wildfire scale prediction using meteorological factors. IEEE Access 7: 176746-176755.

LIU W, WANG S, ZHOU Y, WANG L & ZHANG S. 2010. Spatial distribution patterns analysis of historical forest fires in DaXingAn Mountains of China. In: 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), 3, p. 634-638.

MAINGI JK & HENRY MC. 2007. Factors influencing wildfire occurrence and distribution in eastern Kentucky, USA. Int J Wildland Fire 16(1): 23-33.

MARTÍN Y, ZÚÑIGA-ANTÓN M & MIMBRERO MR. 2019. Modelling temporal variation of fire-occurrence towards the dynamic prediction of human wildfire ignition danger in northeast Spain. Geomat Nat Haz Risk 10(1): 385-411.

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

MARTÍNEZ-FERNÁNDEZ J, CHUVIECO E & KOUTSIAS N. 2013. Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. Nat Hazard Earth Sys 13(2): 311-327.

MASELLI F, ROMANELLI S, BOTTAI L & ZIPOLI G. 2003. Use of NOAA-AVHRR NDVI images for the estimation of dynamic fire risk in Mediterranean areas. Remote Sens Environ 86(2): 187-197.

MERCER DE & PRESTEMON JP. 2005. Comparing production function models for wildfire risk analysis in the wildland - urban interface. Forest Ecol Manag 7(5): 782-795.

MICHAELSEN J, SCHIMEL D S, FRIEDL MA, DAVIS FW & DUBAYAH RC. 1994. Regression Tree Analysis of satellite and terrain data to guide vegetation sampling and surveys. J Veg Sci 5(5): 673-686.

MILLSPAUGH JJ, NIELSON RM, MCDONALD L, MARZLUFF JM, GITZEN R, RITTENHOUSE CD, HUBBARD MW & SHERIFF SL. 2006. Analysis of resource selection using utilization distributions. J Wildlife Manage 70(2): 384-395.

MORGAN P, HARDY CC, SWETNAM TW, ROLLINS MG & LONG DG. 2001. Mapping fire regimes across time and space: Understanding coarse and fine-scale fire patterns. Int J Wildland Fire 10(4): 329-342.

MYERS N, MITTERMEIER RA, MITTERMEIER CG, FONSECA GAB & KENT J. 2000. Biodiversity hotspots for conservation priorities. Nature 403: 853-858.

OLIVEIRA PE & PASSACANTILI MGSB. 2010. Influência antrópica em três ecótonos floresta/campo da Floresta Atlântica do Sudeste do Brasil: análise de micropartículas carbonizadas em solos superficiais. Hoehnea 37(4): 777-789.

OLIVEIRA S, OEHLER F, SAN-MIGUEL-AYANZ J, CAMIA A & PEREIRA JM. 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. Forest Ecol Manage 275: 117-129.

PARITSIS J, VEBLEN TT & HOLZ A. 2015. Positive fire feedbacks contribute to shifts from *Nothofagus pumilio* forests to fire-prone shrublands in Patagonia. J Veg Sci 26(1): 89-101.

PARZEN E. 1962. On Estimation of a Probability Density Function and Mode. Ann Math Stat 33(3): 1065-1076.

PERRY DA. 1998. The scientific basis of forestry. Annu Rev Ecol Syst 29: 435-466.

PRASAD VK, BADARINATH KVS & EATURU A. 2008. Biophysical and anthropogenic controls of forest fires in the Deccan Plateau, India. J Environ Manage 86(1): 1-13.

PRESTEMON JP, HAWBAKER TJ, BOWDEN M, CARPENTER J, BROOKS MT, ABT KL, SUTPHEN R & SCRANTON S. 2013.Wildfire Ignitions: A review of the science and recommendations for empirical modeling. Gen. Tech. Rep. SRS-GTR-171. Asheville, NC: USDA-Forest Service, Southern Research Station. p. 1-20.

ROSENBLATT M. 1956. Remarks on Some Nonparametric Estimates of a Density Function. Ann Math Stat 27(3): 832-837.

ROY DP, BOSCHETTI L, JUSTICE CO & JU J. 2008. The collection 5 MODIS burned area product — Global evaluation by comparison with the MODIS active fi re product. Remote Sens Environ 112(9): 3690-3707.

SANTOS JF, SOARES RV & BATISTA AC. 2006. Perfil dos incêndios florestal no Brasil em áreas protegidas no período de 1998 a 2002. Floresta 36(1): 93-100.

SEAMAN DE & POWELL RA. 1996. An Evaluation of the Accuracy of Kernel Density Estimators for Home Range Analysis. Ecology 77(7): 2075-2085.

SEMERARO T, MASTROLEO G, ARETANO R, FACCHINETTI G, ZURLINI G & PETROSILLO I. 2016. GIS Fuzzy Expert System for the assessment of ecosystems vulnerability to fire in managing Mediterranean natural protected areas. J Environ Manage 168: 94-103.

SILVERMAN BW. 1986. Density estimation for statistics and data analysis. Monographs on Statistics and Applied Probability, London: Chapman and Hall.

SOS MATA ATLÂNTICA. 2019. Relatório Annual 2018. Available in< https://www.sosma.org.br/sobre/relatorios-e-balancos/> Access in: June 30, 2020.

STEFANIDOU M, ATHANASELIS S & SPILIOPOULOU C. 2008. Health impacts of fire smoke inhalation. Inhal Toxicol 20(8): 761-766.

STEINBERG D & COLLA P. 1997. (CART): Classification and regression trees. Software, Salford-System, San Diego, California.

STOLLE F & LAMBIN EF. 2003. Interprovincial and interannual differences in the causes of land-use fires in Sumatra, Indonesia. Environ Conserv 30(4): 375-387.

SYPHARD AD, MASSADA AB, BUTSIC V & KEELEY JE. 2013. Land use planning and wildfire: Development policies influence future probability of housing loss. Plos One 8(8): 1-12.

SYPHARD AD, RADELOFF VC, KEELEY JE, HAWBAKER TJ, CLAYTON MK, STEWART SI & HAMMER RB. 2007. Human influences on California fire regimes. Ecol Appl 17(5): 1388-1402.

SYPHARD AD, RADELOFF VC, KEULER NS, TAYLOR RS, HAWBAKER TJ, STEWART SI & CLAYTON MK. 2008. Predicting spatial patterns of fire on a southern California landscape. Int J Wildland Fire 17(5): 602-613.

TEBALDI ALC, FIEDLER NC, JUVANHOL RS & DIAS HM. 2013. Ações de Prevenção e Combate aos Incêndios Florestais nas Unidades de Conservação Estaduais do Espírito Santo. Floresta e Ambient 20(4): 538-549.

TABARELLI M & GASCON C. 2005. Lessons from fragmentation research: improving management and policy guidelines for biodiversity conservation. Conser Biol 19(3): 734-739.

TEPLEY AJ, VEBLEN TT, PERRY GL, STEWART GH & NAFICY CE. 2016. Positive Feedbacks to Fire-Driven Deforestation Following Human Colonization of the South Island of New Zealand. Ecosystems 19(8): 1325-1344.

VADREVU KP, EATURU A & BADARINATH KVS. 2010. Fire risk evaluation using multicriteria analysis-a case study. Environ Monit Assess 166: 223-239.

VASCONCELOS MP, SILVA S, TOME M, ALVIM M & PEREIRA JC. 2001. Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. Photogramm Eng Rem S 67(1): 73-81.

VEGA-GARCIA C, LEE BS, WOODARD PM & TITUS SJ. 1996. Applying neural network technology to human-caused wildfire occurrence prediction. AI Applications 10(3): 9-18.

VENABLES WN & RIPLEY BD. 2002. Modern Applied Statistics with S-Plus, 4th ed., New York: Springer, 495 p.

WIGTIL G, HAMMER RB, KLINE JD, MOCKRIN M H, STEWART SI, ROPER D & RADELOFF VC. 2016. Places where wildfire potential and social vulnerability coincide in the coterminous United States. Int J Wildland Fire 25: 896-908.

WILLIAMS DR, JAKES PJ, BURNS S, CHENG AS, NELSON KC, STURTEVANT V & SOUTER SG. 2012. Community wildfire protection planning: The importance of framing, scale, and building sustainable capacity. J Forest 110(8): 415-420.

WORTON BJ. 1989. Kernel Methods for Estimating the Utilization Distribution in Home- Range Studies. Ecology 70(1): 164-168.

WOTTON BM, NOCK CA & FLANNIGAN MD. 2010. Forest fire occurrence and climate change in Canada. Int J Wildland Fire 19(3): 253-271.

WU X & KUMAR V. 2009. The Top Ten Algorithms in Data Mining. New York: Chapman and Hall/CRC, 232 p.

WU Z, HE HS, YANG J, LIU Z & LIANG Y. 2014. Relative effects of climatic and local factors on fire occurrence in boreal forest landscapes of northeastern China. Sci Total Environ 493: 472-480.

YANG J, HE HS, SHIFLEY SR & GUSTAFSON EJ. 2007. Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. Forest Sci 53(1): 1-15.

ZHANG Y, LIM S & SHARPLES JJ. 2016. Modelling spatial patterns of wildfire occurrence in South-Eastern Australia. Geomat Nat Haz Risk 7(6): 1800-1815.

ZHANG G, WANG M & LIU K. 2019. Forest fire susceptibility modeling using a convolutional neural network for Yunnan Province of China. Int J Disast Risk Sc 10: 386-403.

SUPPLEMENTARY MATERIAL

Figures S1-S4.

How to cite

JUVANHOL RS, FIEDLER NC, SANTOS AR, PELUZIO TMO, SILVA WB, PINHEIRO CJG & SOUSA HCP. 2023. Use of machine learning as a tool for determining fire management units in the brazilian atlantic forest. An Acad Bras Cienc 95: e20201039. DOI 10.1590/0001-3765202320201039.

Manuscript received on July 2, 2020;
accepted for publication on October 19, 2020

RONIE S. JUVANHOL[1]
https://orcid.org/0000-0002-0040-3382

NILTON CESAR FIEDLER[2]
https://orcid.org/0000-0002-4376-3660

ALEXANDRE R. DOS SANTOS[2]
https://orcid.org/0000-0003-2617-9451

TELMA M.O. PELUZIO[3]
https://orcid.org/0000-0003-0462-9239

WELLINGTON B. DA SILVA[4]
https://orcid.org/0000-0003-2242-7825

CHRISTIANO JORGE G. PINHEIRO[4]
https://orcid.org/0000-0003-2898-8738

HELBECY CRISTINO P. DE SOUSA[1]
https://orcid.org/0000-0002-2683-8827

[1]Federal University of Piaui/UFPI, BR 135, Km 03,
Planalto Horizonte, 64900-000 Bom Jesus, PI, Brazil

[2]Federal University of Espírito Santo/UFES, Postgraduate Programme in Forest Sciences, Av. Governador Lindemberg, 316, Centro, 29550-000 Jerônimo Monteiro, ES, Brazil

[3]Federal Institute of Espírito Santo, Campus Alegre, Rodovia ES 482, Km 47, 29500-000 Alegre, ES, Brazil

[4]Federal University of Espírito Santo/UFES, Department of Rural Engineering, Alto Universitário, s/n, 29500-000 Alegre, ES, Brazil

Correspondence to: **Ronie Silva Juvanhol**
E-mail: roniejuvanhol@ufpi.edu.br

Author Contributions

JUVANHOL RS, Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. FIEDLER NC, SANTOS AR, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – review & editing. PELUZIO TMO, SILVA WB, PINHEIRO CJG, Formal Analysis, Supervision, Writing – review & editing. SOUSA HCP, Visualization, Writing – review & editing.