



Wildfire susceptibility mapping: Deterministic vs. stochastic approaches



Michael Leuenberger^{a,*}, Joana Parente^c, Marj Tonini^b, Mário Gonzalez Pereira^{c,d}, Mikhail Kanevski^b

^a Institute of Statistics, Faculty of Science, University of Neuchâtel, Neuchâtel, 2000, Switzerland

^b Institute of Earth Surface Dynamics (IDYST), Faculty of Geosciences and Environment, University of Lausanne, Lausanne, 1000, Switzerland

^c Centro de Investigação e de Tecnologias Agro-Ambientais e Biológicas (CITAB), Universidade de Trás-os-Montes e Alto Douro, Vila Real, 5000-801, Portugal

^d Instituto Dom Luiz (IDL), Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Edifício C8, Piso 3, 1749-016 Lisboa, Portugal

ARTICLE INFO

Article history:

Received 23 March 2017

Received in revised form

20 December 2017

Accepted 21 December 2017

Available online 6 January 2018

Keywords:

Susceptibility mapping

Wildfires

Random forest

Extreme learning machines

Portugal

ABSTRACT

Wildfire susceptibility is a measure of land propensity for the occurrence of wildfires based on terrain's intrinsic characteristics. In the present study, two stochastic approaches (i.e., extreme learning machine and random forest) for wildfire susceptibility mapping are compared versus a well established deterministic method. The same predisposing variables were combined and used as predictors in all models. The Portuguese region of Dão-Lafões was selected as a pilot site since it presents national average values of fire incidence and a high heterogeneity in land cover and slope. Maps representing the susceptibility of the study area to wildfires were finally elaborated. Two measures were used to compare the different methods, namely the location of the pixels with similar standardized susceptibility and total validation burnt area. Results obtained with the stochastic methods are very alike with the deterministic ones, with the advantage of not depending on a priori knowledge of the phenomenon.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Wildfires are defined as unwanted fires occurring in countryside or rural area and burning forest and wild lands, included abandoned agricultural lands and rural vegetated areas. Wildfires, as undesirable as often uncontrolled events, represent a hazardous and harmful phenomena to people and environment. Natural fires, caused by lightning, appeared on the Earth surface in concomitance with the first plant communities, well before the appearance of humans, and played a key role in plant adaptation and the ecosystems equilibrium (Pausas and Keeley, 2009). Nowadays, the primary cause of wildfires in populated areas is related to the human activities that voluntary (arsonism) or involuntary (accidental or negligent causes) can initiate fire. A recent analysis of fire data from the European Forest Fire Information System shows that over 95% of wildfires are human induced (San-Miguel-Ayanz et al., 2012) and this percentage is even higher in the Mediterranean regions.

Estimating the probability of wildfire occurrence in a certain

area under particular environmental and anthropogenic conditions is a modern tool to support forest protection plans and to reduce fires' consequences, which can also affect the neighbouring or intermingled urban areas. In this context, the implementation of wildfire susceptibility maps and the investigation of the main driving factors inducing wildfires is fundamental. A good review of these factors can be found in (Ganteaume et al., 2013): they included human factors and related variables (such as distance to road or to urban area) as well as environmental factors. More or less sophisticated models have been applied to combine the predisposing variables into a geographic information systems (GIS) (Chuvieco et al., 2010; Chuvieco and Salas, 1996; Bonazountas et al., 2005; Jaiswal et al., 2002). The most reliable analyses applied statistical models to assess the importance of different variables influencing fire occurrences and the obtained results are used to produce the risk maps (Beverly et al., 2009; Soto et al., 2013; Pourtaghi et al., 2015). Recent analyses compared different statistical models for variable selection (Pourghasemi, 2016; Pourghasemi et al., 2016; Pourtaghi et al., 2016; Rodrigues et al., 2014; Eugenio et al., 2016) but most of the studies relied on expert knowledge to pre-select most important drivers or on the results of linear (deterministic) statistical models.

* Corresponding author.

E-mail address: michael.leuenberger@unine.ch (M. Leuenberger).

Portugal is unequivocally the European country most affected by wildfires, due to its favorable climatic conditions, topography and vegetation (Amraoui et al., 2015; Pereira et al., 2013). Investigations of driven factors and the elaborations of wildfires density and risk maps were latterly performed for this highly affected country. Tonini et al. (2017) analysed the spatio-temporal density distribution of these hazardous events in the last decades and produced a 3D graphical output of the results, which highlights areas and frame-periods more affected by wildfires. Nunes et al. (2016) used geographically weighted regression to identify relevant municipal drivers of fires. It results that topography and population density were significant factors in municipal ignitions, while topography and uncultivated land were significant factors in municipal burnt area (BA). Verde and Zêzere (2010) assessed forest fire susceptibility, testing and using variables of strong spatial correlation (i.e. elevation, slope, land cover, rainfall and temperature) and, more recently, Parente and Pereira (2016) adopted this method, updating the selected variables, to map the structural fire risk in the vegetated area of Portugal.

In the present study, the authors refer to the wildfire susceptibility mapping as an estimation of the probability that fire occurs in a specific area without considering a temporal scale, assessed on the basis of predisposing factors related to terrain's intrinsic characteristics. The revised literature misses the use of stochastic models to elaborate accurate susceptibility maps of wildfires, which can be compared with the results obtained by applying deterministic approaches. These latter methods usually assume a priori knowledge of predisposing factors, or they are evaluated by applying linear methods, which implies that every set of variable states is uniquely determined by the parameters used in the model and by the sets of previous states. Therefore, a deterministic model always performs the same way for a given set of initial conditions. Contrary to the deterministic approach, the stochastic methods assume that results obtained by the combination of independent factors (i.e. variables), affecting the investigated phenomenon, can be slightly different due to the randomness of the process. This aspect is particularly useful to model environmental and anthropogenic hazard, which naturally present a complex behaviours and patterns.

Therefore, the objective of the present study is to compare stochastic approaches vs a well established deterministic method for wildfire susceptibility mapping. A first assessment of the susceptibility and hazard wildfire performed for Portugal (Verde and Zêzere, 2010; Parente and Pereira, 2016) is used as benchmarking while extreme learning machine (ELM) and random forest (RF) are the two applied stochastic methods. We restricted our investigation to a pilot area, namely the region of Dão-Lafões, characterized by a high variability and heterogeneity of environmental features and fire incidence similar to the national average, which makes it a good representative of the general characteristics of Continental Portugal.

2. Material: study area and datasets

2.1. Study area

Portugal is the European country more to the southwest, with a Mediterranean type of climate, but suffering of the influence of the Atlantic Ocean that bathes its western and southern coasts (Parente et al., 2016). Mainland Portugal has a total land area of about 90'000 km², which, according to the Corine Land Cover (CLC) 2006 inventory, is predominantly used for agriculture (47%), followed by forests coverings (23%), scrub and/or herbaceous vegetation associations (23%) and open spaces with little or no vegetation (2%) (Pereira et al., 2014).

According to the Planos Regionais de Ordenamento Florestal (PROF), Continental Portugal is divided into 21 PROF regions (Fig. 1). The PROF establish specific rules for the use and exploitation of its forest spaces, in order to ensure sustainable production of all goods and services associated with them (ICNF, 2016).

In the present study, Dão-Lafões region was selected as the case study area for the following reasons: (i) it is located in the Northern half of Portugal, which presents, by far, the highest wildfire incidence (Parente and Pereira, 2016); (ii) this region presents an annual average number of fires and BA very similar to the national average and; (iii) its area is very heterogeneous in terms of topography, land use and vegetation cover (Fig. 2).

2.2. The datasets

Raw data used in this study include: (i) Digital Elevation Model (DEM) derived from the Shuttle Radar Topographic Mission with a resolution of 1 arc-second (DEM-SRTM ~ 25 m), used to compute elevation and slope (Gonçalves and Morgado, 2008); (ii) CLC 2006 inventory, produced by the European Environment Agency, which provides the land use and land cover maps; and, (iii) the National Mapping Burnt Areas (NMBA) implemented by the Institute for the Conservation of Nature and Forests (ICNF, 2016), which provides a detailed description of the shape and the size of the area burnt by fires in each year of occurrence. The data pre- and post-processing, as well as the mapping elaboration, were performed by Quantum GIS free software (QGIS Development Team, 2016).

2.2.1. Topography

Topography, characterized by the altitude, slope and exposure, constitutes one of the most important factors to define the type of the climate of a region such as the average weather conditions and the space-time variability of the climatic elements (e.g., air



Fig. 1. Location of Mainland Portugal and its 21 PROF regions protruding the study region of Dão-Lafões.

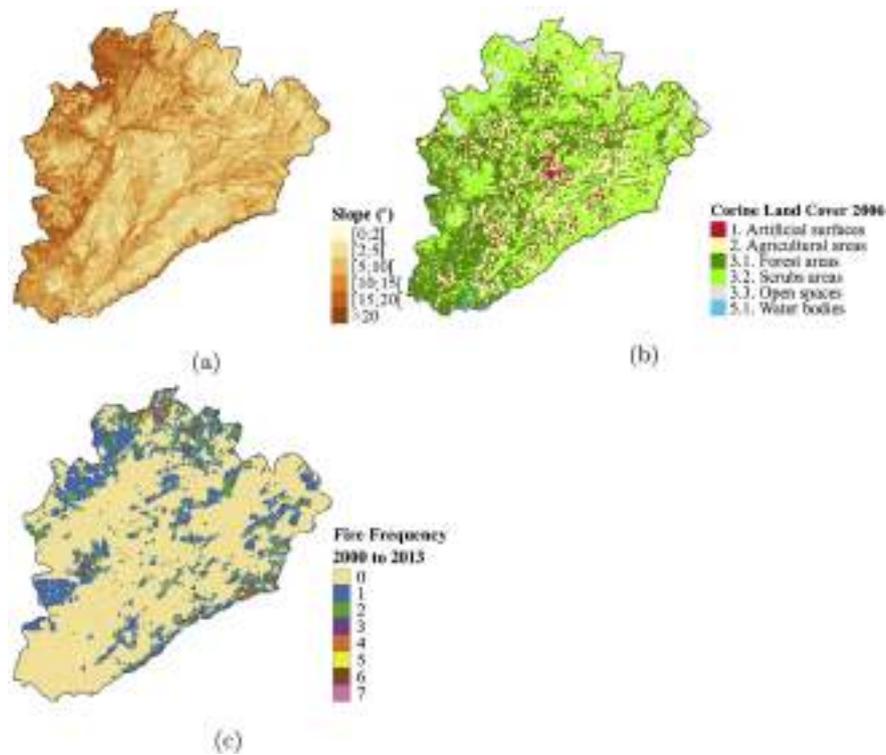


Fig. 2. (a) Slope, (b) land cover according to Corine Land Cover 2006 inventory, and (c) fire frequency for the 2000–2013 period in the Portuguese PROF region of Dão-Lafões.

temperature, precipitation, solar radiation). These factors control the life cycle of the vegetation cover and land use and have a profound influence on the fire incidence (Chuvieco and Congalton, 1989; Freire et al., 2002; Verde and Zêzere, 2010; Parente and Pereira, 2016; Parente et al., 2016). In this study, we considered the slope as the main topographic variable influencing the susceptibility to wildfires in the study area. This value was derived from the DEM and categorized in the same 6 classes used by Verde and Zêzere (2010), namely: 0–2%, 2–5%, 5–10%, 10–15%, 15–20% and >20%.

2.2.2. Land use and vegetation cover

The CLC consists of an inventory of land cover in 44 classes with a minimum map unit of 25 ha for areal phenomena. The main classes are: artificial surfaces, agricultural, forest and semi-natural areas, wetlands and water bodies (Büttner, 2014; Caetano et al., 2009). The 2006 version of CLC was used in the present study (Fig. 2), since this date is in the middle of the investigated period (2000–2013). In investigated region (Dão-Lafões), the different classes of land cover and land use are quite homogeneously distributed within the area. However, it is possible to identify some patterns: higher concentration of forest cover may be found in the southwest and middle-class slopes; agricultural areas are mostly located in the southeast, away from the highest slopes, while scrubs are predominant in the southeast and northwest borders as well as in high slopes.

2.2.3. The fire dataset

The NMBA is an official Portuguese fire dataset based on satellite imagery, acquired once per year at the end of the fire season, and delivered in vector format, as polygons of the BA allowing a detailed description of the location, size and shape of the fire scars, which is fundamental for the present study. This dataset was recently reviewed to correct a minor number of missing values and data

inconsistencies. It contains 17'903 fire events between 2000 and 2013, where 1'114 of which occurred on Dão-Lafões (Parente et al., 2016). In this region, most of the fire incidences are located far in the north and in the southeast (Fig. 2), affecting mostly agricultural areas (10%), scrublands (62%) and open spaces (13%) as well as areas with slopes ranging from 5 to 10° (32%) and 10 to 15° (23%). The location and size of the BA for the investigated period (2000–2013) is represented in Fig. 3 in the form of fire frequency (ff), which is the number of times each pixel burnt over the fourteen years. The year with the highest fire incidence was 2005 (with 11% of the total number of fires and 29% of the total BA in the study period), followed by 2012 (11% of the total number of fires) and 2013 (8% of the total number of fires and 20% of the total BA). Only 18% of total number of pixels burnt at least once and the fire frequency is mostly low or very low (Figs. 2 and 3), with 97% of the total number of burnt pixels (TNBP) with $ff < 3/14$ (namely, 72% of TNBP with $ff = 1/14$, 20% of TNBP with $ff = 2/14$ and 5% of TNBP with $ff = 3/14$).

3. Methodology

Both deterministic and stochastic models for wildfire susceptibility mapping were applied in the present study. The deterministic model, used as benchmark, was developed by Verde and Zêzere (2010) and further adopted and updated by Parente and Pereira (2016). The model includes the computation of fire occurrence probability and favorability scores for each predisposing variable (land cover and slope). Two stochastic methods from the machine learning field were then applied: RF and ELM. Generally speaking, stochastic models account for the uncertainty in modelling processes that have some kind of randomness and, therefore, are useful to represent phenomena with random variability. In the case of machine learning algorithms, the models produce susceptibility maps based on input data (variables) without the need of a priori

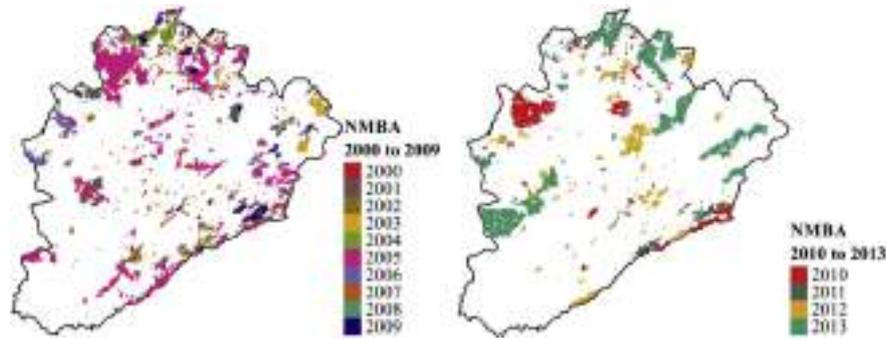


Fig. 3. Annual burnt area polygons in the calibration (2000–2009) and validation (2010–2013) periods in the Portuguese PROF region of Dão-Lafões.

knowledge of the investigated phenomena, but simply learning from experience. Once the model is fitted according to the training data, it allows to generate predictions over the entire study area. In the present study, data were splitted into training (2000–2009) and validation periods (2010–2013): the first was used to fit and calibrate the three models and the second to assess and compare susceptibility maps.

The susceptibility maps were elaborated by means of GIS procedures and organized into 5 classes, in agreement with the Portuguese law (DL, 2006). The classes were defined as in the reference works (Verde and Zêzere, 2010; Parente and Pereira, 2016) using the quintiles of the susceptibility, computed as explained below. The applied methods were assessed by computing the matching, pixel by pixel, between the standardized susceptibility maps obtained for the training period (2000–2009) and the effective BA over the validation period (2010–2013). These values were finally evaluated as a percentage for each susceptibility class.

The next two subsections are devoted to the brief description of the deterministic and stochastic applied models.

3.1. Deterministic method

In Portugal, national authorities, such as Forest Service (ICNF) and the Meteorological Office (*Instituto Português do Mar e da Atmosfera*, IPMA) adopted the wildfire susceptibility map proposed by Verde and Zêzere (2010) which was developed using a deterministic approach and based on just three factors. The susceptible values for each regular unit-area (i.e., pixel) is computed by integrating the favorability scores (Fav) of the two variables (slope and vegetation cover) and the fire probability (fp) as:

$$SP = fp \cdot Fav_{\text{slope}} \cdot Fav_{\text{vegetation}} \quad (1)$$

The favorability scores for each class x ($Fav(x)$) of slope and vegetation cover are computed by:

$$Fav(x) = \frac{NBP(x)}{TNP(x)} \times 100, \quad (2)$$

where $NBP(x)$ is the number of burnt pixels in class x and $TNP(x)$ is the total number of pixels in the class x . The fire probability of each pixel is estimated using the fire database and the classic definition of probability according to:

$$fp = \frac{(\text{the number of times the pixel burned in the study period, in years})}{(\text{duration of the study, in years})} \times 100. \quad (3)$$

It is important to note that, due to the yearly temporal acquisition of the fire database (NMBA), each pixel can only burn once in each year. In addition, due to the multiplicative nature of susceptibility equation, all the null favorability scores were reclassified to one, thus becoming neutral values in the equation. Therefore, the obtained value in each pixel is a consequence of all the possible combinations of the variables found in that pixel.

3.2. Machine learning algorithms

At present, machine learning algorithms are important tools for the analysis, modelling and visualization of environmental data (Kanevski et al., 2009). They have good generalization abilities when modelling high dimensional and complex nonlinear phenomena, are universal modelling methods and many of them have solid roots in statistical learning theory (Hastie et al., 2009). In predictive learning, they focus on modelling the hidden relationship between a set of input and output variables by trying to minimize both the errors and the complexity of the model. After a training procedure, to calibrate the model's parameters, prediction maps of the susceptibility can be computed and displayed with the corresponding uncertainty quantification. In this study, two machine learning algorithms, based on two different concepts, are used for comparison purposes: RF, which is based on decision trees, and ELM, which is based on traditional artificial neural networks. Detailed application of the RF and ELM for environmental data modelling along with the description of consistent methodology are presented in literature (Micheletti et al., 2014; Leuenberger and Kanevski, 2015). Analysis were performed using R free software (R Core Team, 2016). The packages *randomForest* and *elmNN* were employed for RF and ELM respectively.

3.2.1. Random forest

Developed by Breiman (2001), RF is an ensemble machine learning algorithm based on decision trees. It contains two hyperparameters: the number of decision trees generated ($nbtree$), and the number of selected variables for each split node ($nbvar$).

The random forest algorithm first generates $nbtree$ subsets of the training dataset by bootstrapping (i.e. random sampling with replacement). Then, for each subset, it will grow a decision tree by iterating the following rules up to the maximum level (when each final node contains less than 5 data points):

- 1 for each split, the algorithm selects randomly *nbvar* variables,
- 2 according to these *nbvar* variables and the output variable, it computes the Gini index (Hastie et al., 2009) and selects the best variable with the best threshold in order to minimize the error of the prediction.

Once the *nbtree* decision trees have been grown, prediction of new data points is performed by taking the average value of all decision trees (Fig. 4):

$$y_{\text{pred}} = \frac{1}{nbtree} \sum_{i=1}^{nbtree} y_i. \quad (4)$$

This procedure leads to a robust mean value of prediction as well as a measure of uncertainty by considering the standard deviation among all trees.

3.2.2. Extreme learning machine

ELM is based on artificial neural network concept. Following the structure of a single-hidden layer feedforward neural network (SLFN), it connects all input variables to the hidden layer, computes the neurone value and averages all neurons, with optimal weights, to the output layer (Huang et al., 2006; Leuenberger and Kanevski, 2015). More formally, composed of *nbnode* neurons (\tilde{N}) and by using an activation function $g : \mathbb{R} \rightarrow \mathbb{R}$, the ELM network, connecting the inputs (\mathbf{x}_i) to the output (y_i) value, can be written in the following form:

$$\sum_{j=1}^{\tilde{N}} \beta_j g(\mathbf{x}_i \cdot \mathbf{w}_j + b_j) = y_i, \quad (5)$$

where $\mathbf{x}_i \cdot \mathbf{w}_j$ is an inner product between the input (\mathbf{x}_i) and the weight vector (\mathbf{w}_j) which connects the input layer to the *j*th neuron, b_j is the bias of the *j*th neuron, and β is a weight vector connecting the hidden layer to the output layer. In a more compact way, ELM can be written as:

$$H\beta = \mathbf{y}, \quad (6)$$

where $H_{ij} = (u_j)_{j=1, \dots, \tilde{N}} = g(\mathbf{x}_i \cdot \mathbf{w}_j + b_j)$ is the output matrix of the hidden layer (Fig. 5).

According to this notation, ELM algorithm applies the following steps:

1. Randomly generates the input weight \mathbf{w}_j and the bias b_j ;
2. Computes the matrix H ;
3. Computes the output weight $\beta = H^\dagger \mathbf{y}$, where H^\dagger is the Moore-Penrose generalized inverse of matrix H .

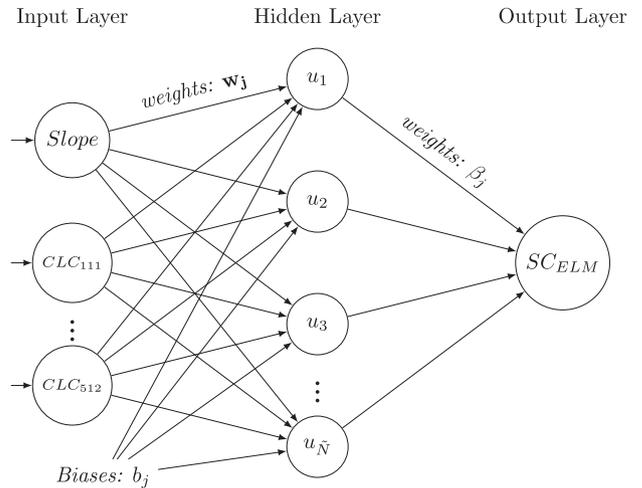


Fig. 5. Structure of the ELM following a single-hidden layer feedforward neural network (SLFN). In this configuration, slope and CLC classes are used as input variables, and SC_{ELM} stands for the susceptible value of the model.

At the end, the only fitted parameter is the number of hidden neuron ($\tilde{N} = nbnode$).

3.2.3. Parameter optimization

In order to optimize the learning process of RF and ELM, different pre-processing steps must be considered. First of all, the CLC classes were converted into 27 dummy variables (one for each class of the CLC variable within the study area). Then, the complete dataset, which is composed of 28 input variables (slope + CLC variables) and 1 output variable (presence or absence of forest fire), was normalized into the [0, 1] interval. This transformation was performed in order to fit the functional range where ELM works in an optimal way (Huang et al., 2006). After that, from the 5'581'522 raster cells covering the study area, 100'000 points (approximately 1.8% of the total area) were randomly selected using a stratified sampling for the construction of the testing subset (Table 1). Namely, 6 strata were used by considering the number of time each pixel burnt (between 0 and 5 times, in this case). This process was reiterated 20 times in order to generate 20 training subsets, but without considering already selected testing points.

The optimization of the *nbtree* and *nbvar* hyper-parameters of RF was performed by using a trial and error process. The choice of this method is justified by the fact that both types of RF hyper-parameters are not highly sensitive to changes and optimized values are close to the default ones. In this study, hyper-parameters of RF were set to 1000 and 9 for *nbtree* and *nbvar*, respectively. For

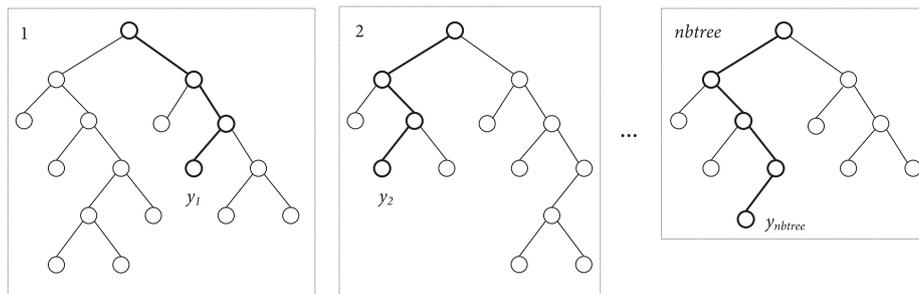


Fig. 4. Structure of RF based on an ensemble of single decision trees.

Table 1

Total number of pixels for the entire study area (All dataset) with the corresponding proportion of pixel which burnt 0, 1, 2, 3, 4 and 5 times. The column *Testing and training subsets* displays the proportion which was used in order to generate the 20 training subsets and the testing subset.

	All dataset	Testing and training subsets
# of pixels which never burnt	4'535'233	81'254
# of pixels which burnt once	856'554	15'346
# of pixels which burnt twice	166'232	2'978
# of pixels which burnt 3 times	20'127	361
# of pixels which burnt 4 times	3'049	55
# of pixels which burnt 5 times	327	6
Total number of pixels	5'581'522	100'000

ELM, the *nbnode* hyper-parameter, which is the number of nodes in the hidden layer, a 5-fold cross-validation approach was performed. Minimum mean squared error (MSE) values obtained for the validation sets were retained, which lead to an optimal number of 40 nodes for this dataset.

Once each machine learning algorithm was fitted, 20 models were built by using the 20 training subsets. Finally, susceptibility maps were generated by averaging the prediction values of the 20 models for the whole study area. In addition to the mean prediction values, standard deviation maps could be extracted from this process and analysed in order to eventually detect areas with high variability in fire susceptibility. A useful by-product of RF algorithm is the variable importance ranking (Breiman, 2001). From an internal evaluation of each variable (based on random shuffling), it computes the percentage of mean square error increase (%IncMSE) by comparing the difference of RF performance when considering both the raw variables and the shuffled variables (Breiman, 2001). As a result, each variable can be ranked according to their %IncMSE score with the following meaning: high %IncMSE score indicates an important contribution in the relationship between input and output variables, while low %IncMSE score (close to 0) indicate that the variable is not a valuable contributor to the model.

4. Results and discussion

The following section 4.1 discusses the selection of CLC and slope variables as the only parameters influencing the wildfire susceptibility in our study area. Then, results and comparisons on the susceptibility maps generated by the three proposed methods are presented in section 4.2. Finally, section 4.3 assesses the different methods by using data from a testing period. Moreover, it presents the variable importance measurement for both the deterministic and the random forest approaches and discusses on the relevance of the obtained results according to the literature in this field.

4.1. Major variables affecting wildfires occurrence in Portugal

The deterministic model was first proposed by Verde and Zêzere (2010), further discussed in Verde (2015), and then updated and used by Parente and Pereira (2016) to map the structural fire risk. This model is based on the combination of geographic variables that do not change much in the short period. This is in line with the wildfire susceptibility, being a measure of the terrain/land propensity for the occurrence of wildfires based on the terrain's intrinsic characteristics (Parente and Pereira, 2016).

Although it is a quite simple model, parsimoniously based on just two variables, it is very robust. Its robustness was recently assessed (Verde, 2015) in respect to the use of single or multiple CLC inventories as well as to rely the calibration and validation on different CLC inventories. The obtained results point to a relative

independence of the model performance in relation to how many or which CLC inventories are used to access the favourability scores. Parente and Pereira (2016) test the impacts of using a high-resolution DEM and, besides mapping the susceptibility with higher spatial accuracy, the obtained patterns were very similar. Obviously, changes in vegetation cover and different fire history can induce different susceptibility patterns due to changes fire probability and vegetation dynamics.

Many researchers have studied the fire processes/mechanisms and tried to identify the underlying factors in Portugal including topography, land use land cover, climate, man-made features, demographic and socio-economic information. For example, Nunes et al. (2016) found that topography, land cover, population density and livestock are significant in both ignition density and BA. Variables such as altitude, slope and land cover help to explain the existence of space-time clusters of fires in Portugal (Parente and Pereira, 2016; Parente et al., 2016; Tonini et al., 2017).

Verde and Zêzere (2010) tested the usefulness of other variables such as altitude, temperature and precipitation in the deterministic model, but they did not find any significant increase in the prediction rates. This may be due to several reasons. First, some of these variables can be proxies of each other. Slope is a measure of the altitude change (Chang and Tsai, 1991; Parente and Pereira, 2016) while altitude regulates the rainfall and temperature (Li et al., 2010; Neteler et al., 2011; Parente and Pereira, 2016). Climate/weather conditions determines the existence, type and state of the vegetation at each location, which means that the information about vegetation cover implicitly considers climate information (Parente and Pereira, 2016). Second, all fires tend to occur associated to high air temperature, low humidity and relatively long periods of drought (Amraoui et al., 2015; Parente and Pereira, 2016; Trigo et al., 2006). Third, vegetation cover can be viewed as a set of different variables instead of just one. For example, to model fire ignition probabilities, Vasconcelos et al. (2001) test the usefulness of CLC related variables such as distance to urban areas, distance to agricultural areas, distance to forests, distance to scrublands, etc., which can be viewed as a different use of vegetation cover. Oliveira et al. (2012) adopted a similar procedure to study the spatial distribution of large fires, considering the proportion of forest area, of scrubs, of agricultural areas, etc. Finally, in a very recent study, Fernandes et al. (2016) identifies fuels and topography as the major determinants of large-size BA in Portugal and in the Western Mediterranean Basin, which is consistent with previous findings on the characterization of wildfires in Portugal (Marques et al., 2011).

Another aspect that must be pointed out in deterministic model is the double use of the BA/fire probability, namely: (i) to compute the favourability scores to rank CLC and slope classes in terms of fire proneness; and, (ii) in the expression of susceptibility, in the form of fire probability in each pixel, i.e., to discriminate where, within the country, each class is more or less affected. In addition, fire probability is also a proxy for the human behaviour since the large majority of the fires are caused by humans (Parente and Pereira, 2016; Verde and Zêzere, 2010).

4.2. Susceptibility maps

Fig. 6 shows the susceptibility maps obtained by applying the three models. In a broad sense, the three models lead to relatively similar maps. The main areas with high/very high and low/very low susceptibility classes are detected and highlighted on similar locations. The very high susceptibility class shows a common pattern for the three models and is mainly located on the North of the region and on the South border.

In order to evaluate the two stochastic models, assuming the

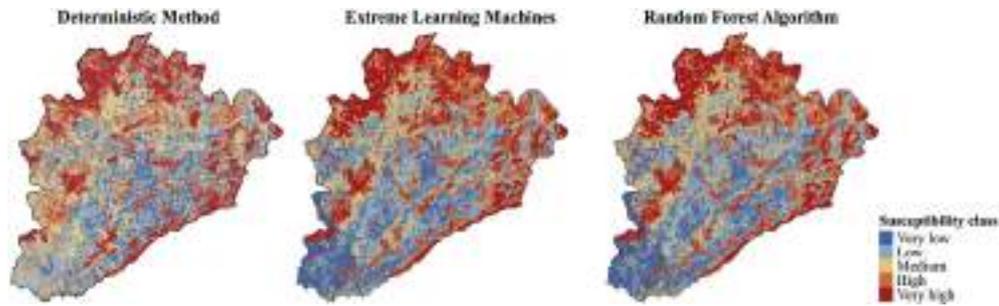


Fig. 6. Susceptibility maps for the three models based on the 5 generated classes, which are very high, high, medium, low and very low.

deterministic one as reference, maps of the differences of susceptibility were generated (Fig. 7). These maps were produced by assigning each class of susceptibility to a unique value (very low = 0, low = 1, medium = 2, high = 3 and very high = 4), and by computing the differences pixel by pixel. These maps are predominantly characterized by light colours (Fig. 7), which means that differences, when they exist, occur between successive classes ($-1 \leq diff \leq 1$). The southwest part of the study area shows an apparent and systematic underestimation of the susceptibility classes for RF and ELM models compared with the deterministic model (Fig. 7a and b). Nevertheless, this difference is not problematic since it concerns essentially of classifying a pixel in the very low class instead of low or medium classes. For the rest of the region, differences between the stochastic and deterministic models are insignificant. Differences between the two machine learning algorithms (RF and ELM) are shown on Fig. 7c: these differences are only slightly present but without significant spatial variations. This result is mainly due to the same pre-processing and similar methodological procedure featuring the two stochastic methods. Moreover, from the machine learning point of view, the use of 100'000 training points contributes to the general stability of both RF and ELM models.

Finally, it is important to note that standard deviation maps (not shown) computed from the 20 RF and ELM models built by using the 20 training subsets to eventually detect areas with high variability in fire susceptibility, reveal, on the contrary, very low variability of both models. In addition to this, a general evaluation of both methods was performed by computing the mean squared error (MSE) on the testing set (which has never been used during the learning process). Unsurprisingly, ELM and RF algorithms show highly similar results with a MSE of 0.1115 for ELM and 0.1117 for RF.

4.3. Methods' assessment

Fig. 8, Table 2 and Table 3 show the proportion of BA within each

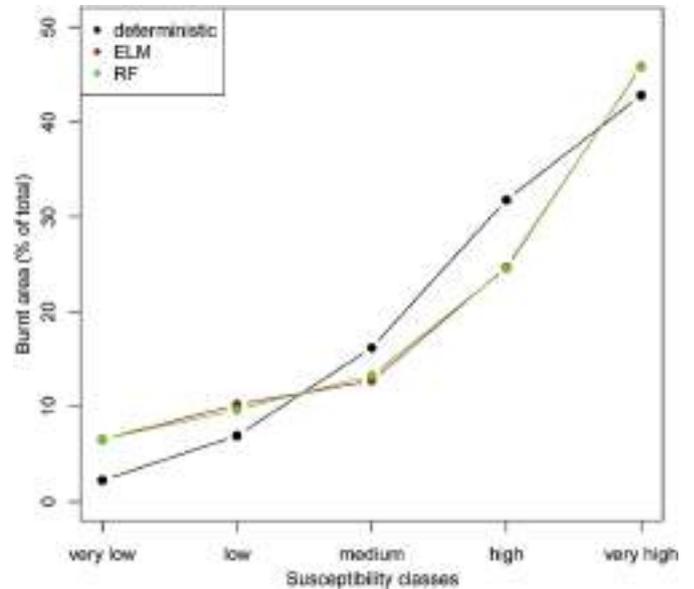


Fig. 8. Proportion of total BA explained by each of the three models in each susceptibility class.

Table 2

Ratio between the size of each class and the proportion of BA for the testing period for the deterministic approach.

Susceptibility classes	ratio	Nb Pixels	%
Very high	33.1	360'161	42.8
High	22.6	267'356	31.8
Medium	10.4	136'410	16.2
Low	6.6	58'392	6.9
Very low	1.7	18'875	2.3
Total	15.1	841'194	100.0

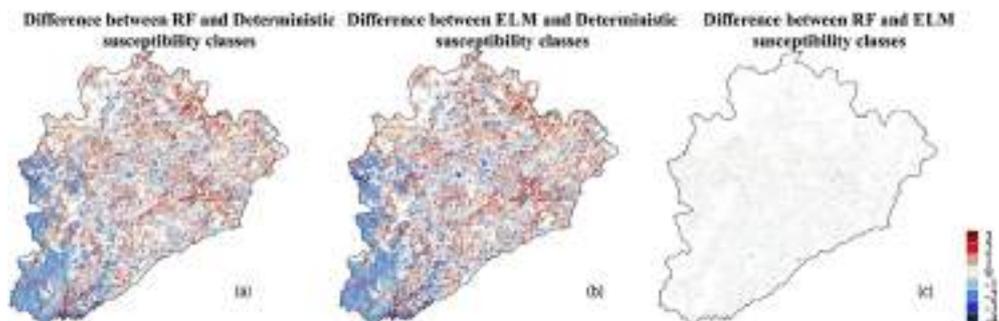


Fig. 7. Differences of the susceptibility classes between the three methods. For each susceptibility class of Fig. 6, a value is assigned (very low = 0, low = 1, medium = 2, high = 3 and very high = 4) and the differences between classes for each model were computed.

Table 3

Ratio between the size of each class and the proportion of BA for the testing period for ELM and RF.

Susceptibility classes	ELM			RF		
	ratio	Nb Pixels	%	ratio	Nb Pixels	%
Very high	34.5	385'519	45.8	34.6	386'051	45.9
High	18.6	207'640	24.7	18.5	206'570	24.6
Medium	9.6	107'178	12.8	10	111'970	13.3
Low	7.7	85'877	10.2	7.3	81'211	9.7
Very low	4.9	54'980	6.5	5	55'392	6.5
Total	15.1	841'194	100.0	15.1	841'194	100.0

susceptibility class obtained for each method and assessed for the testing period (2010–2013). Moreover, the ratio between the size of each class and the proportion of BA were also computed. By considering the deterministic model as the reference, the ELM and RF susceptibility maps reveal proportions of BA close to the benchmark model. Differences in the percentage of total BA in each susceptibility class is always less than 7.2%. Apparently, the percentage of total BA in the two first susceptibility classes (very low and low) is higher for ELM and RF (approximately 4%), but in the medium and high susceptibility classes this value is higher for the deterministic approach (about 3%–7%). For the last susceptibility class (very high), both ELM and RF algorithms show a percentage of total BA higher than the deterministic one (3% higher).

As it was already mentioned in section 4.1, the susceptibility maps generated by ELM and RF are very similar. As shown in Fig. 8, Tables 2 and 3, the maximum difference between the percentage of total BA in each susceptibility class for both methods is 0.5% in low and medium classes, and almost zero in the others classes. Generally, and by considering the three approaches, the obtained results are promising in the sense that less than 20% of the total BA of the testing period was classified as very low or low susceptibility (by summing the very low and low scores). This evaluation over the testing period allows to validate the proposed new approach in this field through machine learning algorithms, and to compare the stochastic and deterministic approaches on non-used dataset.

The RF algorithm allows an internal evaluation of each input variable, which leads to a variable importance ranking (Breiman, 2001). This last result constitutes a significant added value to the understanding of the phenomenon. In Fig. 9, the top 9 variables for RF are listed by decreasing order of their respective %IncMSE score.

The first six land cover variables (CLC324, CLC322, CLC334, CLC333, CLC312 and CLC321) represent the variables that most contribute to model and explain the observed variance (higher %IncMSE score). These correspond to: transitional woodland-shrub, moors and heathland, burnt areas, sparsely vegetated areas, coniferous forest and natural grasslands. This short list is dominated by scrub and/or herbaceous vegetation associations (level 32

of CLC classes) followed by open spaces with little or no vegetation (level 33) and forests (level 31). These results are in accordance with fire selectivity studies performed for Portugal where fire selectivity is generally higher for scrublands, pine stands and eucalyptus plantations than for evergreen oak woodlands, annual and rainfed crops and agroforestry lands (Barros and Pereira, 2014). Similar findings were recently obtained for fire proneness studies, also performed for Portugal (Moreira et al., 2009; Silva et al., 2009). In general, agricultural areas are excluded from this list because it includes well managed arable lands (both irrigated and non irrigated), permanent crops (vineyards, olive groves, fruit trees and berry plantations and even pastures). However, heterogeneous agricultural areas (CLC level 24), especially those corresponding to complex cultivation patterns with significant areas of natural vegetation, present higher relative importance in RF stochastic models. The slope is one of the most important factors of fire spread, acting on different aspects of the fuel combustion (Rothermel, 1972). However, per se, i.e. without the other aspects of the fire environment/controls usually conceptualized in fire triangles (e.g. Whitlock et al., 2010), the slope is not able to independently determine the terrain/land propensity for the occurrence or spread of a wildfire. For example, terrain parcel with high slope can be free from vegetation. Therefore, it is not surprising that the ranking of the most important variables is dominated by the land cover variables with 9 classes in the top 10 variables.

On Table 4, the top 6 variables for both random forest and deterministic methods are shown. For comparison purposes the favorability score of the deterministic model (computed based on eq. (2)) are retained. As highlighted, 5 of the 6 top variables selected by random forest are also among the most important variables of the deterministic model even if with a different order. This fact underlines that, in spite of the differences between the methods (random forest being able to detect non-linear relationship), the matching between the most relevant variables is highly satisfactory and validates the use of the new approach based on machine learning algorithm.

The apparent greater importance of conifers (CLC312) in relation to the mixed (CLC313) and broadleaf forest/hardwoods (CLC311) for the RF (Fig. 9) is also worth noting for two reasons: (1) these variables present the same relative importance in both methods; and, (2) it is in good agreement with previous studies for vegetation fire proneness performed for Portugal (Silva and Harrison, 2010; Pereira et al., 2014). In fact, the increase in conifer tree component tends to increase the difficulties to control the fire (Rowe and Scotter, 1973), BA and fire proneness (Moreira et al., 2009; Silva et al., 2009; Silva and Harrison, 2010) and fire risk in WUIs (Lampin-Maillet et al., 2010).

5. Conclusion

In the present paper, susceptibility maps of wildfires obtained by applying stochastic methods, namely Random Forest and Extreme Learning Machine, were compared with the correspondent map elaborated by applying a validated standard deterministic method, here considered as a benchmark. The study was performed for the Dão-Lafões region of Portugal, which is a representative region of a country highly prone to wildfires. The variables, implemented into the model, considered as favorable factors for wildfires, are the slope, the land use and vegetation covers, provided by the Corine Land Cover 2006 inventory. The official dataset of the national mapping BA was considered to train (2000–2009 period) and test (2010–2013 period) the models. Comparison of the obtained results clearly suggests that the two stochastic models perform in an equal manner in terms of

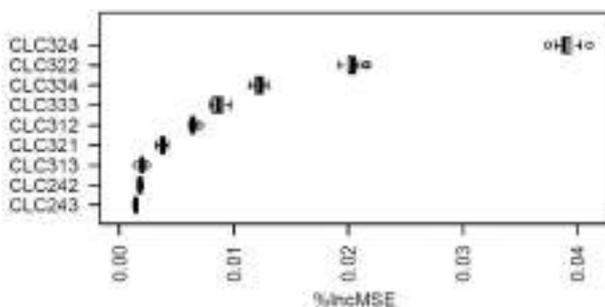


Fig. 9. Variable importance computed with random forest algorithm over 20 runs. The 9 top variables are displayed with the corresponding % increase of mean square error.

Table 4
Variable importance for the top 6 feature for the random forest and the corresponding rank in deterministic model. Favorability scores computed in the deterministic model are used for ranking the variables. CLC classes are identified by the level code.

Variables description	Random Forest		Deterministic approach	
	CLC classes	%IncMSE	Rank position	Favorability score
Traditional woodland-shrub	324	0.03906	6	48.45520
Moors and heathland	322	0.02036	3	68.22487
Burnt areas	334	0.01219	1	95.71093
Sparsely vegetated areas	333	0.00862	2	83.87631
Coniferous forest	312	0.00646	15	4.86921
Natural grasslands	321	0.00378	4	59.39358

susceptibility areas and classes as well as that these results are broadly consistent with susceptibility maps obtained with the benchmarking model. The main benefit of using stochastic models is that these approaches are data driven, meaning that they do not need a priori knowledge of the process. Moreover, random forest directly provides the measurement of the importance of each variable. On this respect, the RF and the deterministic models present similar top variable importance ranking. Results of the present analysis are encouraging for further applications of stochastic models to elaborate susceptibility maps considering more variables and larger areas.

6. Software and data availability

The following software and data were used to perform the analysis presented in this paper:

- **QGIS** (QGIS Development Team, 2016), an open source geospatial software, was mainly used for the pre- and post-processing and the elaboration of maps.
- **R language** (R Core Team, 2016) is an open source statistical software. It was used with the packages *randomForest* and *elmNN* for computing the random forest and the extreme learning machine algorithms.
- **Digital Elevation Model (DEM)** derived from the Shuttle Radar Topographic Mission (STRM - NASA) was used to compute the slope.
- **Corine Land Cover (CLC 2006)** is an inventory provided by the European Environment Agency. It was used in order to extract the land use and land cover map.
- **National Mapping Burnt Areas (NMBA)** is an official Portuguese fire dataset and provides a detailed description of the shape and the size of BA. It was provided by the Institute for the Conservation of Nature and Forests (ICNF, 2016).

Acknowledgements

Michael Leuenberger wants to thank the Institute of Earth Surface Dynamics (University of Lausanne), where he works on this paper during his PhD. This work was supported by: (i) the Herbetta Foundation of the University of Lausanne, under the project 2016-2-E-15; (ii) the FIREXTR project, PTDC/ATPGE0/0462/2014; (iii) the project Interact - Integrative Research in Environment, Agro-Chain and Technology, NORTE-01-0145-FEDER-000017, research line BEST, co-financed by FEDER/NORTE 2020; and, (iv) European Investment Funds by FEDER/COMPETE/POCI - Operacional Competitiveness and Internationalization Programme, under the project POCI-01-0145-FEDER-006958 and National Funds by FCT - Portuguese Foundation for Science and Technology, under the project UID/AGR/04033/2013. We are especially grateful to ICNF for providing the fire data and to João Pereira for the final spelling and grammar review of the manuscript.

References

- Amraoui, M., Pereira, M.G., DaCamara, C.C., Calado, T.J., 2015. Atmospheric conditions associated with extreme fire activity in the Western Mediterranean region. *Sci. Total Environ.* 524, 32–39. <https://doi.org/10.1016/j.scitotenv.2015.04.032>.
- Barros, A.M., Pereira, J.M., 2014. Wildfire selectivity for land cover type: does size matter? *Plos One* 9. <https://doi.org/10.1371/journal.pone.0084760>.
- Beverly, J.L., Herd, E.P.K., Conner, J.C.R., 2009. Modeling fire susceptibility in west central Alberta, Canada. *For. Ecol. Manag.* 258, 1465–1478. <https://doi.org/10.1016/j.foreco.2009.06.052>.
- Bonazountas, M., Kallidromitou, D., Kassomenos, P.A., Passas, N., 2005. Forest fire risk analysis. *Hum. Ecol. Risk Assess.* 11, 617–626. <https://doi.org/10.1080/10807030590949717>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Büttner, G., 2014. CORINE land cover and land cover change products. In: Manakos, I., Braun, M. (Eds.), *Land Use and Land Cover Mapping in Europe: Practices & Trends*. Springer Netherlands, Dordrecht, pp. 55–74. https://doi.org/10.1007/978-94-007-7969-3_5 (chapter 5).
- Caetano, M., Nunes, V., Nunes, A., 2009. CORINE Land Cover 2006 for Continental Portugal. Instituto Geográfico Português.
- Chang, K., Tsai, B., 1991. The effect of DEM resolution on slope and aspect mapping. *Cartogr. Geogr. Inf. Syst.* 18, 69–77. <https://doi.org/10.1559/152304091783805626>.
- Chuvienco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M.P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J.R., Herrera, M.A., Zamora, R., 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. *Ecol. Model.* 221, 46–58. <https://doi.org/10.1016/j.ecomodel.2008.11.017>.
- Chuvienco, E., Congalton, R.G., 1989. Application of remote sensing and geographic information systems to forest fire hazard mapping. *Rem. Sens. Environ.* 29, 147–159. [https://doi.org/10.1016/0034-4257\(89\)90023-0](https://doi.org/10.1016/0034-4257(89)90023-0).
- Chuvienco, E., Salas, J., 1996. Mapping the spatial distribution of forest fire danger using GIS. *Int. J. Geogr. Inf. Syst.* 10, 333–345. <https://doi.org/10.1080/02693799608902082>.
- DL, 2006. Decreto-Lei no. 124/2006 de 28 de Junho. *Diário Da República - I SERIE-A*.
- Eugenio, F.C., dos Santos, A.R., Fiedler, N.C., Ribeiro, G.A., da Silva, A.G., dos Santos, A.B., Paneto, G.G., Schettino, V.R., 2016. Applying GIS to develop a model for forest fire risk: a case study in Espírito Santo, Brazil. *J. Environ. Manag.* 173, 65–71. <https://doi.org/10.1016/j.jenvman.2016.02.021>.
- Fernandes, P.M., Monteiro-Henriques, T., Guiomar, N., Loureiro, C., Barros, A.M., 2016. Bottom-up variables govern large-fire size in Portugal. *Ecosystems* 19, 1362–1375. <https://doi.org/10.1007/s10021-016-0010-2>.
- Freire, S., Carrão, H., Caetano, M.R., 2002. Produção de cartografia de risco de incêndio florestal com recurso a imagens de satélite e dados auxiliares. IGP, Lisboa.
- 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. Manag.* 51, 651–662. <https://doi.org/10.1007/s00267-012-9961-z>.
- Gonçalves, J.A., Morgado, A.M., 2008. Use of the SRTM DEM as a geo-referencing tool by elevation matching. *The international archives of the photogrammetry. Remote Sensing and Spatial Information Sciences* 37, 879–883.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning: data mining, inference and prediction*, 2 ed. Springer <http://www-stat.stanford.edu/~tibs/ElemStatLearn/>.
- Huang, G.B., Zhu, Q.Y., Siew, C.K., 2006. Extreme learning machine: theory and applications. *Neurocomputing* 70, 489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>.
- ICNF, 2016. O que são os PROF? URL <http://www.icnf.pt/portal/florestas/profs/obj>.
- Jaiswal, R.K., Mukherjee, S., Raju, K.D., Saxena, R., 2002. Forest fire risk zone mapping from satellite imagery and GIS. *Int. J. Appl. Earth Obs. Geoinf.* 4, 1–10. [https://doi.org/10.1016/S0303-2434\(02\)00006-5](https://doi.org/10.1016/S0303-2434(02)00006-5).
- Kanevski, M., Pozdnoukhov, A., Timonin, V., 2009. *Machine Learning for Spatial Environmental Data*. EPFL Press, Lausanne, Switzerland.
- Lampin-Maillet, C., Mantzavelas, A., Galiana, L., Jappiot, M., Long, M., Herrero, G., Karlsson, O., Iossifina, A., Thalia, L., Thanassis, P., 2010. Wildland urban

- interfaces, fire behaviour and vulnerability: characterization, mapping and assessment. In: Silva, J.S., Rego, F., Fernandes, P., Rigolot, E. (Eds.), *Towards Integrated Fire Management - Outcomes of the European Project Fire Paradox*. European Forest Institute, pp. 71–92 (chapter 3.3).
- Leuenberger, M., Kanevski, M., 2015. Extreme Learning Machines for spatial environmental data. *Comput. Geosci.* 85, 64–73. <https://doi.org/10.1016/j.cageo.2015.06.020>.
- Li, S., Zhao, Z., Miaomiao, X., Wang, Y., 2010. Investigating spatial non-stationary and scale-dependent relationships between urban surface temperature and environmental factors using geographically weighted regression. *Environ. Model. Software* 25, 1789–1800. <https://doi.org/10.1016/j.envsoft.2010.06.011>.
- Marques, S., Borges, J.G., Garcia-Gonzalo, J., Moreira, F., Carreiras, J.M.B., Oliveira, M.M., Cantarina, A., Botequim, B., 2011. Characterization of wildfires in Portugal. *Eur. J. For. Res.* 130, 775–784. <https://doi.org/10.1007/s10342-010-0470-4>.
- Micheletti, N., Foresti, L., Robert, S., Leuenberger, M., Pedrazzini, A., Jaboyedoff, M., Kanevski, M., 2014. Machine learning feature selection methods for landslide susceptibility mapping. *Math. Geosci.* 46, 33–57. <https://doi.org/10.1007/s11004-013-9511-0>.
- Moreira, F., Vaz, P., Catry, F., Silva, J.S., 2009. Regional variation in wildfire susceptibility of land-cover types in Portugal: implications for landscape management to minimize fire hazard. *Int. J. Wildland Fire* 18, 563–574. <https://doi.org/10.1071/WF07098>.
- Neteler, M., Rocchini, D., Castellani, C., Rizzoli, A., 2011. Terra and Aqua satellites track tiger mosquito invasion: modelling the potential distribution of *Aedes albopictus* in north-eastern Italy. *Int. J. Health Geogr.* 10 <https://doi.org/10.1186/1476-072X-10-49>.
- Nunes, A.N., Lourenço, L., Meira, A.C.C., 2016. Exploring spatial patterns and drivers of forest fires in Portugal (1980–2014). *Sci. Total Environ.* 573, 1190–1202. <https://doi.org/10.1016/j.scitotenv.2016.03.121>.
- Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., Pereira, J.M., 2012. Modeling spatial patterns of fire occurrence in mediterranean europe using multiple regression and random forest. *For. Ecol. Manag.* 275, 117–129. <https://doi.org/10.1016/j.foreco.2012.03.003>.
- Parente, J., Pereira, M.G., 2016. Structural fire risk: the case of Portugal. *Sci. Total Environ.* 573, 883–893. <https://doi.org/10.1016/j.scitotenv.2016.08.164>.
- Parente, J., Pereira, M.G., Tonini, M., 2016. Space-time clustering analysis of wildfires: the influence of dataset characteristics, fire prevention policy decisions, weather and climate. *Sci. Total Environ.* 559, 151–165. <https://doi.org/10.1016/j.scitotenv.2016.03.129>.
- Pausas, J.G., Keeley, J.E., 2009. A burning story: the role of fire in the history of life. *Bioscience* 59, 593–601. <https://doi.org/10.1525/bio.2009.59.7.10>.
- Pereira, M.G., Aranha, J., Amraoui, M., 2014. Land cover fire proneness in Europe. *Forest Systems* 23, 598–610. <https://doi.org/10.5424/fs/2014233-06115>.
- Pereira, M.G., Calado, T.J., DaCamara, C.C., Calheiros, T., 2013. Effects of regional climate change on rural fires in Portugal. *Clim. Res.* 57, 187–200. <https://doi.org/10.3354/cr01176>.
- Pourghasemi, H.R., 2016. GIS-based forest fire susceptibility mapping in Iran: a comparison between evidential belief function and binary logistic regression models. *Scand. J. For. Res.* 31, 80–98. <https://doi.org/10.1080/02827581.2015.1052750>.
- Pourghasemi, H.R., Beheshtirad, M., Pradhan, B., 2016. A comparative assessment of prediction capabilities of modified analytical hierarchy process (M-AHP) and Mamdani fuzzy logic models using Netcad-GIS for forest fire susceptibility mapping. *Geomatics. Nat. Hazards Rev.* 7, 861–885. <https://doi.org/10.1080/19475705.2014.984247>.
- Pourtaghi, Z.S., Pourghasemi, H.R., Aretano, R., Semeraro, T., 2016. Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. *Ecol. Indic.* 64, 72–84. <https://doi.org/10.1016/j.ecolind.2015.12.030>.
- Pourtaghi, Z.S., Pourghasemi, H.R., Rossi, M., 2015. Forest fire susceptibility mapping in the Minudasht forests, Golestan province, Iran. *Environ. Earth Sci.* 73, 1515–1533. <https://doi.org/10.1007/s12665-014-3502-4>.
- QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial Foundation. <http://qgis.osgeo.org>.
- R Core Team, 2016. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rodrigues, M., de la Riva, J., Fotheringham, S., 2014. Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Appl. Geogr.* 48, 52–63. <https://doi.org/10.1016/j.apgeog.2014.01.011>.
- Rothermel, R.C., 1972. *A Mathematical Model for Predicting Fire Spread in Wildland Fuels*.
- Rowe, J.S., Scotter, G.W., 1973. Fire in the boreal forest. *Quat. Res.* 3, 444–464. [https://doi.org/10.1016/0033-5894\(73\)90008-2](https://doi.org/10.1016/0033-5894(73)90008-2).
- San-Miguel-Ayanz, J., Schulte, E., Schmuck, G., Camia, A., Strobl, P., Libertà, G., Giovando, C., Boca, R., Sedano, F., Kempeneers, P., McInerney, D., Withmore, C., Santos de Oliveira, S., Rodrigues, M., Durrant, T., Corti, P., Oehler, F., Vilar, L., Amatulli, G., 2012. Comprehensive monitoring of wildfires in europe: the european forest fire information system (EFFIS). In: Tiefenbacher, J. (Ed.), *Approaches to Managing Disaster - Assessing Hazards, Emergencies and Disaster Impacts*. InTech, pp. 87–108. <https://doi.org/10.5772/28441> (chapter 5).
- Silva, J.S., Harrison, S.P., 2010. Humans, climate and land cover as controls on european fire regimes. In: Silva, J.S., Rego, F., Fernandes, P., Rigolot, E. (Eds.), *Towards Integrated Fire Management - Outcomes of the European Project Fire Paradox*. European Forest Institute, pp. 49–59 (chapter 3.1).
- Silva, J.S., Moreira, F., Vaz, P., Catry, F., Godinho-Ferreira, P., 2009. Assessing the relative fire proneness of different forest types in Portugal. *Plant Biosystems. Int. J. Dealing with all Aspects of Plant Biology* 143, 597–609. <https://doi.org/10.1080/11263500903233250>.
- Soto, M.E.C., Molina-Martínez, J.R., Rodríguez y Silva, F., Alvear, G.H.J., 2013. A territorial fire vulnerability model for Mediterranean ecosystems in South America. *Ecol. Inf.* 13, 106–113. <https://doi.org/10.1016/j.ecoinf.2012.06.004>.
- Tonini, M., Pereira, M.G., Parente, J., Orozco, C.V., 2017. Evolution of forest fires in Portugal: from spatio-temporal point events to smoothed density maps. *Nat. Hazards* 85, 1489–1510. <https://doi.org/10.1007/s11069-016-2637-x>.
- Trigo, R.M., Pereira, J.M., Pereira, M.G., Mota, B., Calado, T.J., Dacamara, C.C., Santo, F.E., 2006. Atmospheric conditions associated with the exceptional fire season of 2003 in Portugal. *Int. J. Climatol.* 26, 1741–1757. <https://doi.org/10.1002/joc.1333>.
- Vasconcelos, M.P., Silva, S., Tome, M., Alvim, M., Pereira, J.C., 2001. Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. *Photogramm. Eng. Rem. Sens.* 67, 73–81.
- Verde, J.C., 2015. *Wildfire Susceptibility Modelling in Mainland Portugal*.
- Verde, J.C., Zézere, J.L., 2010. Assessment and validation of wildfire susceptibility and hazard in Portugal. *Natural Hazards and Earth System Sciences* 10, 485–497. <https://doi.org/10.5194/nhess-10-485-2010>.
- Whitlock, C., Higuera, P.E., McWethy, D.B., Briles, C.E., 2010. Paleocological perspectives on fire ecology: revisiting the fire-regime concept. *Open Ecol. J.* 3, 6–23. <https://doi.org/10.2174/1874213001003020006>.