A data mining approach for global burned area mapping

Rubén Ramo[a,b,*], Mariano García[a,b], Daniel Rodríguez[c], Emilio Chuvieco[a,b]

[a] Department of Geology, Geography and Environment, University of Alcalá, Colegios 2, 28801 Alcalá de Henares, Spain
[b] Enviromental Remote Sensing Research Group, Department of Geology, Geography and the Environment, University of Alcalá, Colegios 2, 28801 Alcalá de Henares, Spain
[c] Department of Computer Science, University of Alcalá, Ctra. de Barcelona Km 33.6, 28871 Alcalá de Henares, Spain

ABSTRACT

Global burned are algorithms provide valuable information for climate modellers since fire disturbance is responsible of a significant part of the emissions and their related impact on humans. The aim of this work is to explore how four different classification algorithms, widely used in remote sensing, such as Random Forest (RF), Support Vector Machine (SVM), Neural Networks (NN) and a well-known decision tree algorithm (C5.0), for classifying burned areas at global scale through a data mining methodology using 2008 MODIS data. A training database consisting of burned and unburned pixels was created from 130 Landsat scenes. The resulting database was highly unbalanced with the burned class representing less than one percent of the total. Therefore, the ability of the algorithms to cope with this problem was evaluated.

Attribute selection was performed using three filters to remove potential noise and to reduce the dimensionality of the data: Random Forest, entropy-based filter, and logistic regression. Eight out of fifty-two attributes were selected, most of them related to the temporal difference of the reflectance of the bands. Models were trained using an 80% of the database following a ten-fold approach to reduce possible overfitting and to select the optimum parameters.

Finally, the performance of the algorithms was evaluated over six different regions using official statistics where they were available and benchmark burned area products, namely MCD45 (V5.1) and MCD64 (V6). Compared to official statistics, the best agreement was obtained by MCD64 (OE = 0.15, CE = 0.29) followed by RF (OE = 0.27, CE = 0.21). For the remaining three areas (Angola, Sudan and South Africa), RF (OE = 0.47, CE = 0.45) yielded the best results when compared to the reference data. NN and SVM showed the worst performance with omission and commission error reaching 0.81 and 0.17 respectively. SVM and NN showed higher sensitivity to unbalanced datasets, as in the case of burned area, with a clear bias towards the majority class. On the other hand, tree based algorithms are more robust to this issue given their own mechanisms to deal with big and unbalanced databases.

1. Introduction

Wildland fires are one of the most important disturbances in the Earth system, affecting the balance of greenhouse gases (van der Werf et al., 2010), vegetation distribution and society (Goldammer et al., 2008; Kloster et al., 2012; Schoennagel et al., 2009). Wildland fires are considered an Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS) (2004); Hollmann et al., 2013) and has, therefore, been selected by the European Spatial Agency (ESA) as one of the ECV included in the Climate Change Initiative (CCI) program (Hollmann et al., 2013).

Burned area (BA) detection is an active research topic which has been studied over a variety of ecosystems. Many studies have shown the ability of high resolution sensors to map burned areas at local scale using high and medium resolution images (Dragozi et al., 2014; Mitri and Gitas, 2013). Nevertheless, to analyze global vegetation dynamics (Mouillot et al., 2014) or greenhouse gas emissions estimation (Leeuwen et al., 2013), global coverage is needed. In this framework, the most used products are those that use MODIS (Moderate-Resolution Imaging Spectroradiometer) images, such as MCD45 (Roy et al., 2005) or MCD64 (Giglio et al., 2013) products. In addition to these data, there are others BA products developed by different European projects in the last decade such as L3JRC (Tansey et al., 2008), Globcarbon (Plummer et al., 2005) based on SPOT-VEGETATION, or the Fire_cci product (Alonso-Canas and Chuvieco, 2015; Chuvieco et al., 2016) based on MERIS (Medium-Spectral Resolution Imaging Spectrometer). Given the high variety of the burning conditions (i.e. vegetation type, biomass consumption, time prevalence), most of the global BA products relies in

---

the use of regional thresholds to discern between burned and unburned areas (Alonso-Canas and Chuvieco, 2015; Giglio et al., 2013; Plummer et al., 2005; Tansey et al., 2008), but none of them has been yet developed using machine learning algorithms, particularly using a global training dataset.

Data mining, defined as the computing process of discovering patterns and relationship from large dataset through the use of machine learning, statistics and database systems (Fayyad et al., 1996), has experienced an increase of popularity in the remote sensing field because of its capability to extract patterns from apparently unstructured data. For instance, it has been successfully applied to map natural disasters (Barnes et al., 2007; Goswami et al., 2016; Traore et al., 2017), land cover classification (DeFries and Chan, 2000; Zhou et al., 2013) or change detection (Boulila et al., 2011; Hussain et al., 2013). It has also been applied in fire applications such as forest fire prediction (Cheng and Wang, 2008) or to map burned areas (Özbayoğlu and Bozer, 2012; Quintano et al., 2011).

One of the advantages of train global models is that after the training phase, the classification become fully automatic without the need of further calibrations or regional adaptations (Ramo and Chuvieco, 2017). However, the main difficulties of this approach are the necessity of generating a training database that includes the great variability of burned conditions, and the generation of balanced error rate models that classify burned area without overfitting or bias to the majority (or minority) class, obtaining similar error rates results among different regions.

The main objective of this study was to compare the capacity of four well-known machine learning algorithms, namely random forests (RF), support vector machine (SVM), artificial neural networks (ANN) and decision trees (C5.0), to map burned areas at global scale using a data mining approach. The algorithms were applied over six different regions (Australia, Angola, California, South Africa and Sudan) and the results validated in two ways. First, the performance was evaluated by leaving 20% of the training database for independent validation. Second, comparing the BA information yielded by the algorithms with existing official statistics (Australia, Canada and California), and two well-known BA products namely, MCD64 and MCD45.

2. Materials and methods

The proposed methodology consisted of several steps involving the training database compilation, attribute selection, algorithm training and evaluation, image classification and perimeter comparison. The flowchart of the applied methodology is presented in Fig. 1 to facilitate its interpretation.

2.1. Burned Area perimeters

To create the training dataset, the burned area perimeters from the Fire_cci project (http://www.esa-fire-cci.org/ last accessed April 2018) were used. This dataset has been previously used to validate global BA products (Padilla et al., 2015) such as MCD64 (Giglio et al., 2013), MCD45 (Roy et al., 2005) or the Fire_cci product (Alonso-Canas and Chuvieco, 2015). The Fire_cci validation dataset follows a global statistically designed sample (Padilla et al., 2014), thus the training sites were selected using a stratified random sampling were the strata were defined based on the proportion of burned area extracted from the Global Fire Emissions Database (GFED) (Giglio et al., 2013) and the Olson biomes reclassified in 7 categories based on their similarities and fire behavior (e.g. deserts, Tundra and Mangroves were merged in one class). Thus for each biome the proportion of burned area was computed and those with ≥80% of the area burned were grouped into the high burned area, and those with < 80% into the low burned area class, respectively. The Fire_cci validation dataset is composed of 130 Landsat pairs from 2008 (see Fig. 2) covering 1.58 million of km^2 from which 31,578 km^2 correspond to burned area. Burned areas include: Rainfed

cropland (10.10%), mixed forest closed to open > 15% (10.63%), broad-leaved deciduous open 15–40% (5.45%), need-leaved evergreen closed to open > 15% (8.54%), shrubland (14.42%), grassland (16.16%), sparse vegetation (tree, shrub, herbaceous cover > 15%), and vegetation regularly flooded (5.13%).

2.2. MODIS data

The main source of information is the MCD43A4 (v6). This product was developed using Terra and Aqua observations to correct for the BRDF effect (Schaaf et al., 2002). The MCD43A4 has 500 m spatial resolution and includes the spectral information of seven different bands, Red (B1), Near-infrared (NIR, B2), Blue (B3), Green (B4) and three bands in the shortwave infrared region (SWIR, B5-B7). In addition to these bands, several spectral indices were computed to enhance the BA discrimination (Table 1).

2.3. Ancillary data

In addition to the information provided by the spectral bands and indices, information coming from hotspots (HS) was included. Thermal anomalies information has been extensively used for burned area detection because it provides higher contrast between burned and unburned pixels in comparison with other wavelength regions (Alonso-Canas and Chuvieco, 2015; Giglio et al., 2013). Hence, the MODIS MCD14ML (Version 5.1) product, which provides daily global coverage of hotspot with 1 km spatial resolution, was used. Using this data a distance matrix between each pixel to the closest HS was performed and included as an attribute.

Additionally, we included auxiliary data to adapt the model to regional environmental conditions of burned areas. In this case, we used the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) (Danielson and Gesch, 2011) from which the slope and aspect were computed. This information is useful for BA classification since it is related to the fire behaviour and the physical properties of the land. Land cover (LC) information also provides valuable data for BA mapping as it allows for characterizing the fire signal before and after the fire (Moreira et al., 2009) therefore, Land Cover CCI product was used.

Another important factor for BA mapping is related to the ecosystem variation. The condition of burned and the carbon footprint prevalence in the post-fire image is dependent on climate and vegetation type. In order to characterize this factor, we used the Olson biomes (Olson et al., 2001) which divide the world into 16 regions considering their geology, climate, and evolutionary history. Finally, we included the continental regions defined in the Global Fire Emission Database (GFED) that have been developed taking into account how the fire behaves (Giglio et al., 2013) and hence it can help to characterize the burned signal.

2.4. Training database

The database used for this study comprised the spectral and ancillary information previously described for two categories, namely burned and unburned pixels. Regarding the burned area, the database also included information of the burned proportion of the pixel and the date of the burned. The proportion of burned was extracted by overlapping the Landsat perimeters to the MODIS images. The HS was also used to assign the day of burned to each perimeter from the closest HS.

Our approach to map burned area was also based on a multitemporal analysis, therefore, we extracted the MODIS reflectance values for each band from an image acquired prior to the fire (t1) and another one after the fire (t2). For burned pixels, we constrained the search of post-fire images to a period between 2 and 12 days after the day of burned to avoid smoke plumes and clouds. Pixels with no valid observations in this period were rejected from the database. On the other hand, the search of pre-fire information was also constrained to a period of 1 to 10 days. For non-burned pixels, the t1 was set to the
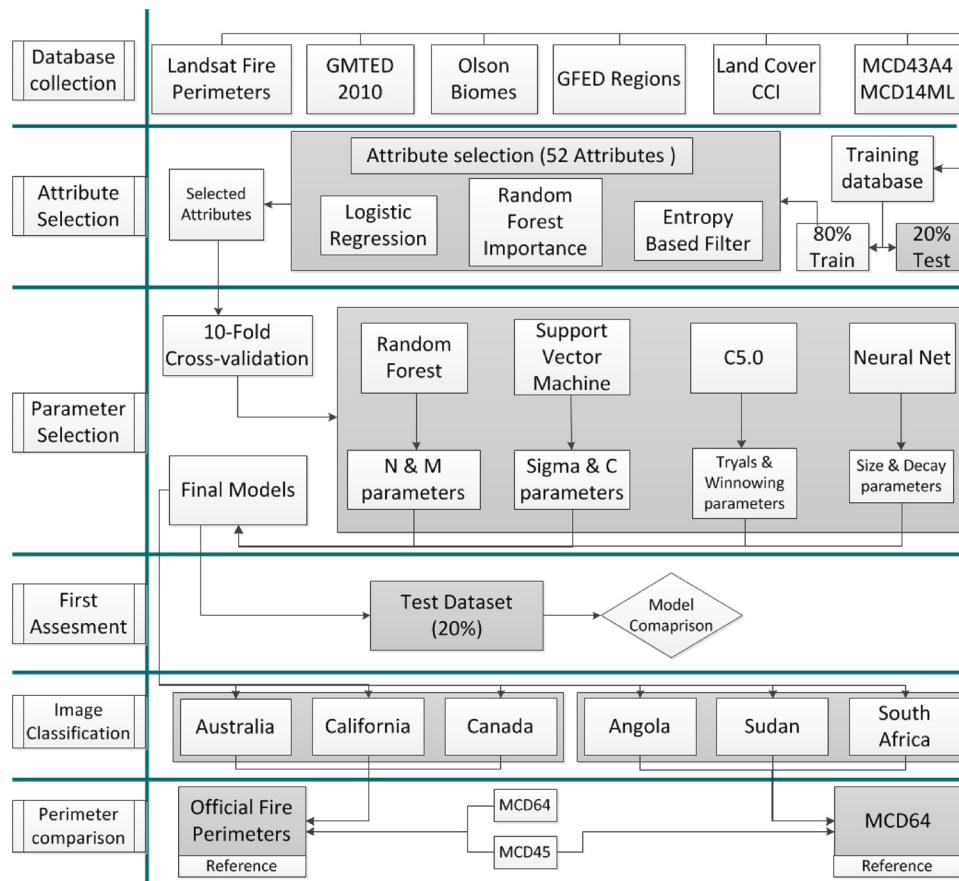
Fig. 1. Methodology flowchart.

prefire Landsat acquisition whereas t2 was set to the median day be-tween the pre and postfire Landsat images. We also included the 10 days threshold in the case there are not available good observations.

After pixel reflectance was extracted, the spectral indices and their differences were calculated for t1 and t2. Subsequently the information contained by the ancillary data (i.e. biome or elevation) was included into the database. Finally, the distance, in meters, between each pixel to

the closest HS occurring between the pre- and post-fire images, were included.

To carry out this process more than 15,000 MODIS images were needed because a Landsat frame can be located in several MODIS tiles. Likewise, it also needed to cover the whole Landsat time gap including the 10 days threshold. In order to reduce the potential noise caused by anomalies in the reflectance or changes produced by an incorrect day of
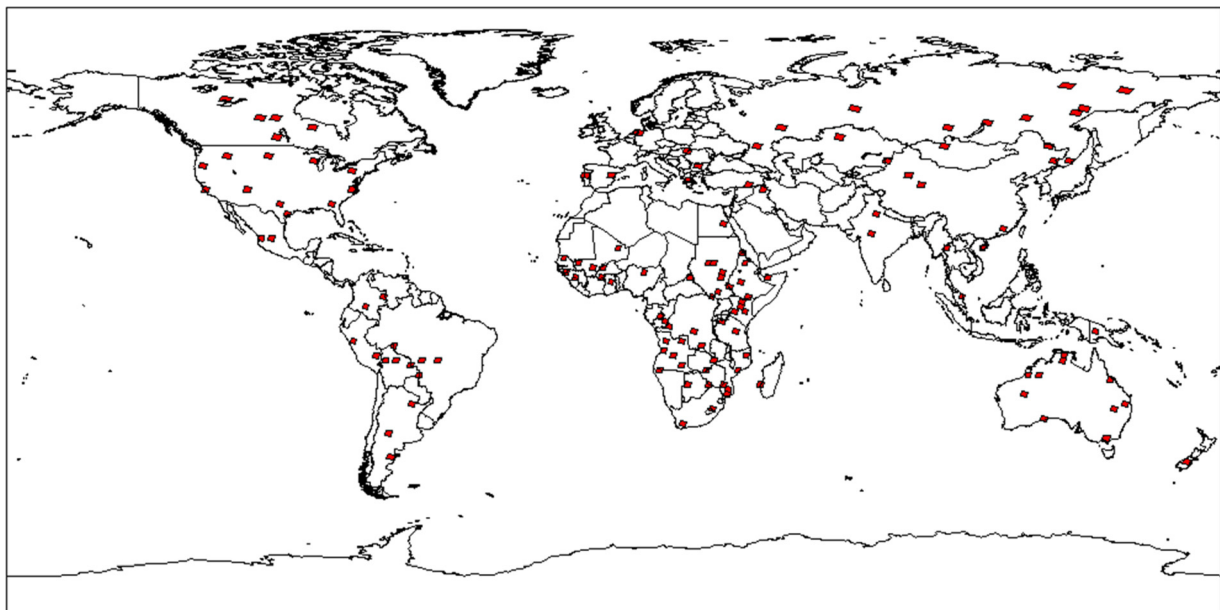

Fig. 2. Distribution of training areas.

Table 1
Indices performed to training database elaboration.

Index	Formula	Parameters	Reference	BA application
NIR Diference	$NIR_{Dif} = \rho_{NIRt1} - \rho_{NIRt2}$	ρ_{NIRt1} = Pre fire NIR band ρ_{NIRt2} = Post fire NIR band	(Alonso-Canas and Chuvieco, 2015)	(Alonso-Canas and Chuvieco, 2015; Giglio et al., 2013)
Soil Adjusted Vegetation Index (SAVI)	$SAVI = \frac{\rho_{NIR} - \rho_{Red}}{(\rho_{NIR} + \rho_{Red} + L)}(1 + L)$	ρ_{NIR} = NIR band ρ_{Red} = Red band L was set to 0.5	(Huete, 1988)	(Chuvieco et al., 2002; Garcia and Chuvieco, 2004)
Global Environmental Monitoring Index (GEMI)	$GEMI = \eta(1 - 0.25\eta) - \frac{(\rho_{Red} - 0.125)}{1 - \rho_R}$ $\eta = \frac{2(\rho_{NIR}^2 - \rho_{Red}^2) + 1.5\rho_{NIR} + 0.5\rho_{Red}}{\rho_{Red} + \rho_{NIR} + 0.5}$	Same meaning than SAVI	(Pinty and Verstraete, 1992)	(Alonso-Canas and Chuvieco, 2015; Smith et al., 2007)
Normalized Burn Ratio (NBR)	$NBR = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}}$	ρ_{NIR} = NIR band ρ_{SWIR} = SWIR band (band 7)	(Garcia and Caselles, 1991)	(Brewer et al., 2005; Loboda et al., 2007; Mallinis and Koutsias, 2012; Rogan and Franklin, 2001)
Normalized Difference Water Index (NDWI)	$NDWI = \frac{\rho_{SWIR} - \rho_{NIR}}{\rho_{NIR} + \rho_{SWIR}}$	ρ_{NIR} = NIR band ρ_{SWIR} = SWIR band (For MODIS band 5 and band 6)	(Gao, 1996)	(Chuvieco et al., 2006; Stroppiana et al., 2003)
Visible Atmospherically Resistant Index (VARI)	$VARI = \frac{\rho_{Green} - \rho_{Red}}{\rho_{Green} + \rho_{Red} - \rho_{Blue}}$	ρ_{Green} = Green band ρ_{Blue} = Blue band ρ_{Red} = Red band	(Gitelson et al., 2002)	(Schneider et al., 2008)
Enhanced Vegetation Index (EVI)	$EVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + 6 \cdot \rho_{Red} - 7.5 \cdot \rho_{Blue} + 1}$	ρ_{Red} = Red band ρ_{NIR} = NIR band ρ_{Blue} = Blue band	(Huete et al., 2002)	(Jin et al., 2012; Quintano et al., 2011; Wittenberg et al., 2007)
Mid-Infrared Burnt Index (MIRBI)	$MIRBI = 10 \cdot \rho_{SWIR3} - 9.8S \cdot \rho_{SWIR2} + 2$	ρ_{SWIR2} = SWIR band (MODIS band 6) ρ_{SWIR3} = SWIR band (MODIS Band 7)	(Trigg and Flasse, 2001)	(Bastarrika et al., 2014; Smith et al., 2007)
Shortwave Angle Slope Index (SASI)	$\beta_{SWIR1} = cos^{-1}\left[\frac{a^2 + b^2 - c^2}{2*a*b}\right]$ $Slope = SWIR2 - NIR$ $SASI = \beta_{SWIR1} * Slope$	a, b and c are the Euclidean distances between NIR and SWIR (MODIS bands 5 and 6)	(Palacios-Orueta et al., 2006)	Not yet tested on burned area.
Angle at NIR (ANIR)	$ANIR\,(\alpha_{SWIR}) = \left[\frac{a^2 + b^2 - c^2}{2*a*b}\right]$	Same meaning as SASI	(Khanna et al., 2007)	Not yet tested on burned area.

burned, we introduced three different filters:

- All burned pixels that show an increase of NIR between t1 and t2 were removed from the database. Since there is a decay in the reflectance in this region short after the fire (Alonso-Canas and Chuvieco, 2015), pixels showing $NIR_{t2} > NIR_{t1}$ were considered as false detections and therefore, removed from the database.
- Unburned pixels having an HS in a 3,000 m radius in the previous 90 days of the Landsat prefire image were removed to avoid possible noise coming from previously burned pixels.
- Pixels burned less than 80% were removed from the analysis in order to avoid mixed signals. Only those that are clearly burned have been included because the classification of burned pixels with mixed spectral signals is very difficult to implement at global scale.

The final database used in this work consisted of 48,464 burned and 6,293,106 unburned instances (pixels) with 52 attributes including the MODIS bands (14 features), the auxiliary variables (7 features), the spectral indices (21 features), and the difference between pre and post-fire indices (10 features). In this frame, a dataset is considered unbalanced when the distribution of the categories is not equal. In this case burned class is highly unbalanced (< 1%) and it produce a bias towards the majority class.

2.5. Attribute selection

Feature selection (FS) aims at reducing a dataset dimensionality by removing irrelevant and redundant attributes while keeping important ones. FS brings several benefits when applying data mining. First, it reduces the risk of overfitting, i.e. lack of model generalization, since models are generally simpler. Second, improves processing time allowing us to explore a larger number of machine learning algorithms. Finally, FS can avoid the collection or calculation of unnecessary attributes for the models used.

The first method used was based on a RF (Breiman, 2001) approach.

RF is an iterative algorithm that creates a group of decision trees; the final result of the classification is based on the results of the whole group of trees. It has two parameters: N, which is the number of decision trees that contain the forest and M, the number of attributes used to perform a decision tree. Although generally applied for classification and regression problems, it can also be used for reducing the dimensionality of datasets. In particular, we used the Boruta implementation of RF (Kursa et al., 2010). Boruta uses RF to estimate the importance of each attribute taking into account that the low correlation between trees is not completely true (Kursa et al., 2010). The Boruta algorithm selects those attributes which have higher importance than the median.

The second method used was based on a logistic regression, which has been previously used to estimate the contribution of the explanatory variables in fire occurrence models (Martinez-Fernandez et al., 2013; Nieto et al., 2012), and to select attributes in a fire ignition probability model (Jurdao et al., 2012). The algorithm train several models making a different combination of attributes and evaluating their performance using a cross-validation approach. The final selection is provided by the model which shows less error rate.

The last method used for feature selection is the Entropy-based filter. Entropy is defined as a measure of the level of impurity of an attribute (Mitchell, 1997). The entropy-based filter is a ranking FS algorithm where each attribute can be measured independently to determine their usefulness for discriminating burned and unburned pixels. The algorithm returns a ranking with a score of importance for each attribute selecting those that have more importance than the median.

Our final attribute selection was based on merging the results of the three methods applied. Thus, only those attributes that were selected by all three approaches were kept for training the algorithms.

2.6. Machine learning algorithms

Burned area classification was based on four common machine learning algorithms that have shown good performance in classifying

Table 2
Parameter grid used for tuning each different classification algorithm.

Algorithm	Parameter one	Parameter two
Random Forest (12 models)	N: 600, 1000, 1200	M: 2, 4, 6, 8
Support Vector Machine (6 models)	Sigma: 0.01, 0.1	C: 10, 50, 100
C5 (8 models)	Winnow: True, False	Trials: 20, 30, 50, 100
Neural Net (15 models)	Size: 1, 3, 5, 10, 15	Decay: 0.1, 0.5, 1

remotely sensed data. Machine learning algorithms do not make any assumptions about the distribution of the data and they have shown better performance in classifying remotely sensed data than traditional parametric approaches (Deng et al., 2016; García et al., 2017; Naidoo et al., 2012).

The first algorithm tested was a random forest (RF). This is one of the most popular classifiers used in remote sensing nowadays (Belgiu and Drăguţ, 2016). Since RF is an ensemble method, it is possible to estimate the probability of a class based on the proportion of trees that categorize a pixel into a class. Despite its popularity in remote sensing studies, its use in forest fire applications has been limited to fire occurrence prediction (Oliveira et al., 2012) and fire regime characterization (Aldersley et al., 2011; Archibald et al., 2009) but its application for burned area classification is limited (Ramo and Chuvieco, 2017).

The second method applied was a support vector machine (SVM), which is a supervised machine learning algorithm developed by Vapnik (2013) based on the statistical learning theory. It attempts to fit an optimal separating hyperplane to the training samples in a multi-dimensional feature space using the idea of structural risk minimization. In this study, the kernel used was a radial basis function (RBF), which has been widely applied in remote sensing applications (Garcia et al., 2011; García et al., 2017; Waske and Benediktsson, 2007) and also for burned area classification (Dragozi et al., 2014; Pereira et al., 2017). The performance of SVM using a RBF is controlled by two parameters, g and C. The parameter g determines the influence of a single feature on the algorithm and C represents the penalty and controls the trade-off between errors and model complexity.

The third method used was based on Artificial Neural Networks (ANN) (Picton, 2000), which consist of a number of basic units or nodes (neurons) distributed in different layers including an input, an output and one or several hidden layers, which determine the architecture of the ANN. The neurons in the input layer receive the attributes of the instances and the neurons of the output layer the class to which each pixel is assigned. The NN algorithm needs to specify two hyper-parameters: size and decay. Size is the number of units in the hidden layer and is related to the number of connections between neurons, and decay is the regularization parameter used to avoid over-fitting. ANNs have been widely used for remote sensing applications but their use for BA has been limited (Al-Rawi et al., 2001; Gómez and Martín, 2011). The wide variety of uses of NN and the ability to handle large dataset make this algorithm a good candidate for classifying burned area (Chu and Guo, 2013; Petropoulos et al., 2010) at global scale.

The last algorithm applied was the C5.0 algorithm (Quinlan, 1993), which uses the gain ratio criterion (Mitchell, 1997) to select the most relevant attribute at every node of the tree during its construction (winnow parameter). The other user-defined parameter is the trials, which enables a boosting procedure where several trees are generated. This classification approach let the model estimate the probability of burned instead of a direct class assignation. The C 5.0 algorithm has been widely applied for remote sensing applications (Igor Klein et al., 2012; Lawrence and Moran, 2015) but has been barely applied for BA detection.

2.7. Model training and algorithm validation

The performance of the different machine learning algorithms

proposed depends on the values of the different parameters required for each method. In order to find the best combination of parameters to ensure their highest performance and avoid overfitting issues, we trained the algorithms following a k-fold cross validation (CV) approach. First, the dataset was divided into training (80%) and testing (20%) groups following a stratified random sample, where the strata were land cover and burned area. Parameter tuning was carried out using the training data (80%), leaving out the rest of the data to validate the models (covered in the next section). The k-fold approach (Bengio and Grandvalet, 2004) implies dividing the training database into k folds, 10 in our study, and use k-1 parts of the data to train the model and use the left out part to evaluate it. After that, a new model is trained permuting the test and the training data parts. Each trained model is assessed using a standard parameter of accuracy allowing the detection of possible overfitting in the case that the results between the ten models calibrated from the cross-validation are significantly different. We used the Area Under the Curve (AUC) (López et al., 2014) to assess the accuracy of the models for each parameter combination.

This process is repeated for each algorithm and parameter combination (Table 2) to find the one that better detects burned pixels.

Each parameter combination has an associated AUC, and the final model will be the one with higher AUC value. The main advantage of this approach is it can also be applied when the database is unbalanced (López et al., 2014). Finally, we calibrated the model using the entire training dataset with the best parameter combination, and their performance was evaluated using confusion matrices based on the remaining 20% of the database that had not been used in the training process. The confusion matrices allowed for the construction of a set of statistics to assess the algorithms quality, namely omission and commission errors and relative bias as an indicator of the error balance (Padilla et al., 2015).

2.8. Global burned area mapping

We attempted to map burned areas globally for 2008. To do so, pairs of MODIS images separated by a 3 days gap were selected. The first image corresponding to the pre-fire situation corresponded to the DOY (day of the year) 1 and the post-fire image to DOY 3. After deriving the attributes selected by our FS approach, the 4 machine learning algorithms were applied. Subsequently, the temporal window embracing the pre- and post-fire images was rolled one day and the classification algorithms were applied to the new dataset. This process was repeated until the end of the year. When the process was over, all binary BA images were merged in a yearly composite. A modal 3×3 filter was applied to give more spatial consistency removing boundary errors and to avoid the characteristic salt and pepper effect of pixel-based classifiers. This provides a smoother image while reducing omission and commission errors (Makido et al., 2007; Yang and Liu, 2005).

RF and C5.0 have the possibility of estimating the probability of the class assigned to the pixel, i.e. the probability of burned. Since the database was highly unbalanced, this allowed us to investigate the minimum probability associated to each pixel to classify it as burned. Thus, a sensitivity analysis was performed to test the effect of burned probability of the pixel, from 10% to 90%, to commission, omission and error balance.

2.9. Comparison with existing BA information

In addition to the validation metrics described above, the accuracy of the models was evaluated by comparing the annual composites with the official fire database of three different areas (Fig. 3). These areas are representative samples of three different fire regimes, namely boreal, tropical and temperate.

The tropical fire regime was represented by our Australian test site, which covered 1,192,585 km^2. The North Australian Fire Information System (www.firenorth.org.au/nafi2/, last accessed April 2018)
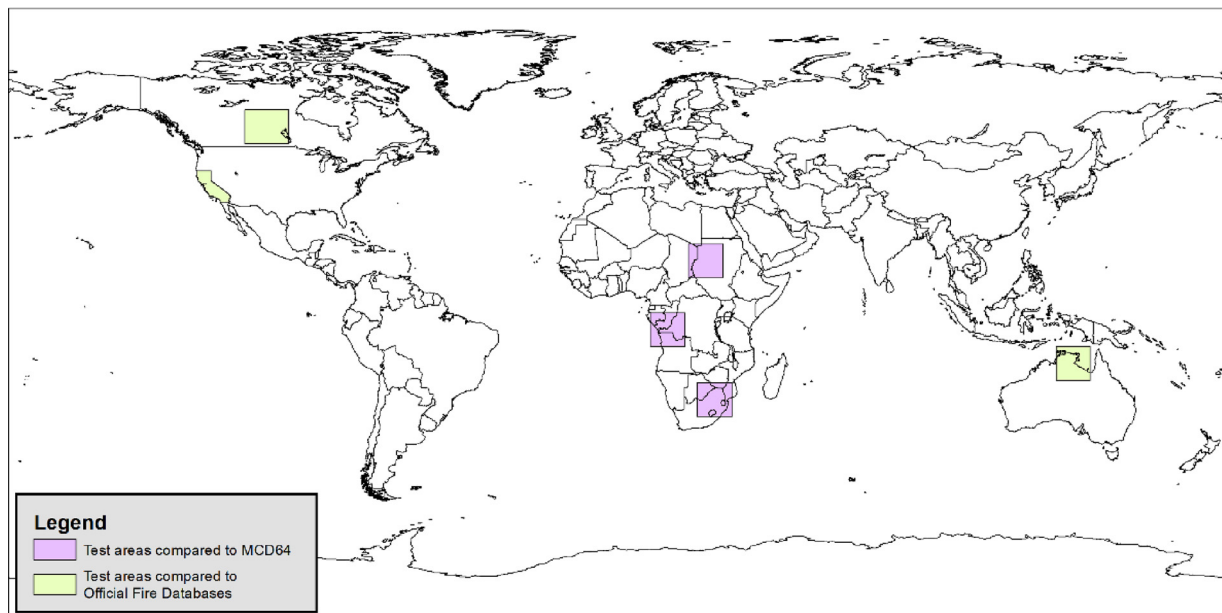
Fig. 3. Location of the test areas.

developed by the Darwin Center for bushfires research, which provides burned area maps generated through multitemporal analysis of 250 m MODIS images using segmentation and visual interpretation, was used as a reference.

The temperate fire regime was represented by our California site. Burned perimeters were obtained from the Fire and Resource Assessment Program (FRAP) webpage (www.frap.fire.ca.gov, last accessed April 2018), which makes available the fire perimeters of the entire state of California (409,719 km^2). These perimeters are produced by several entities like CAL FIRE/FRAP, the USDA Forest Service, Region 5 Remote Sensing Lab, the Bureau of Land Management, and the National Park Service.

For the Boreal fire regime, a region of 926,167 km^2 was selected covering the provinces of Manitoba and Saskatchewan in the central part of Canada. Burned perimeters were downloaded from the Canadian Wildland Fire Information System (www.cwfis.cfs.nrcan.gc.ca/ha/nfdb/, last accessed April 2018). This fire database has been previously used for fire regime characterization in different studies (Burton et al., 2009; Parisien et al., 2006).

With regards Africa, it presents the highest fire incidence in terms of number of fires and amount of burned area of the world (Giglio et al., 2013). Therefore, it is very important to develop an algorithm that performs well over this continent. Nevertheless, there is a lack of validation data to test the performance of the models in this region. Hence, we compared the results of the classifications with the MCD64 product, which currently is the most used source of BA information by climate and atmospheric modellers (Giglio et al., 2013). The models were evaluated in three different areas of Africa with different fire regimes and vegetation types. The first area corresponds to Angola (MODIS tile h19v09), which represents tropical and Subtropical Moist Broadleaf Forests with a very high fire frequency. The other two areas are located in Sudan (h20v07) and South Africa (h20v11).

The total area covered by these six sites is 6,128,741 km^2 and the majority land cover for this areas are: sparse vegetation (29%), croplands (15%), shrubland (14%), tree cover (13%), grassland (12%), seasonality flooded areas (10%) and mixed cropland with natural vegetation (6%).

The annual burned area composites were compared using cross tabulation analysis with the fire datasets. In addition, we extended this analysis to the MCD64 (v6) and MCD45 (v5.1) products to compare the performance of our algorithms with two standard burned area products

except for Africa, where MCD64 was used as a reference. Although this comparison cannot be considered a proper validation exercise, it can provide a first assessment of the methodology. Through the confusion matrix generated in the cross-tabulation analysis, the omission, commission and relative bias were obtained.

3. Results

3.1. Attribute selection

Fig. 4 shows the results of the three FS methods applied. The final attribute selection is composed by 8 different features. Four of the selected variables are based on the difference between the post and the pre-fire signals; one corresponds to the pre-fire situation, one to the post-fire scenario, one is based on distance to hotspots and one is a categorical variable related to the fire behaviour, the GFED region.

RF and the entropy-based filter selected 30 and 27 attributes respectively, whereas the logistic regression was more restrictive selecting only 12 features out of 52 available. It can be seen that different approaches result in different feature importance, therefore, by merging the three approaches we reduce the influence of the FS method used ensuring that only the most important variables are used.

3.2. Model training

The impact of the values of the parameters of each algorithm in the accuracy of the models is shown in Fig. 5.

The NN algorithm was the most affected by the parameter variation, although the change in AUC was approximately 6%, showing that the model accuracy did not change much with the parameters. The number of units in the hidden layer (size) had a positive impact on the accuracy and a medium value of weight variation (decay) released a better error rate. The final model had a size 15 (hidden units) and a decay of 0.5.

On the other hand, the C5.0 algorithm showed little influence of the parameters on its accuracy. The best results were achieved when the winnow parameter was set to false, and the trials to one hundred. The RF models also show similar trends as C5.0. There was a negligible variation in the AUC values between all models (less than 1%) showing better results when the number of attributes (M parameter) was higher. The number of trees did not manifest a significant effect on the accuracy. The final model was trained using 600 trees and six attributes as *n*
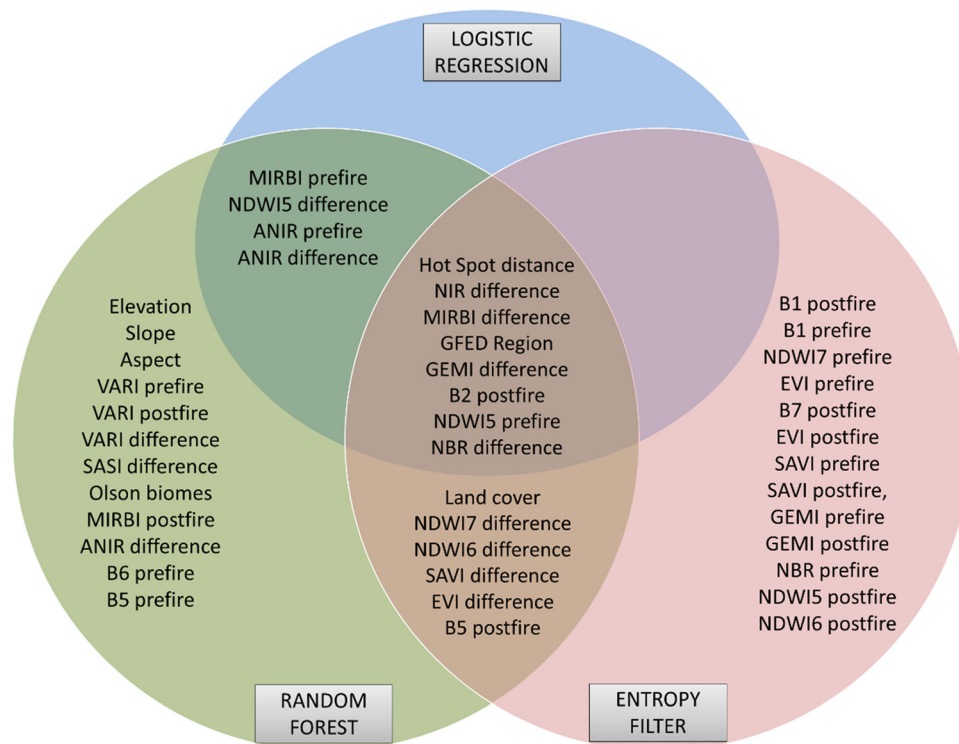
Fig. 4. Attribute selection results represented in a Venn-diagram. The intersection of the three circles represents the common variables selected by the three methods.

and *m* parameters respectively.

For SVM, the width of the kernel (Sigma) had a larger influence on the model performance than the penalty parameter, although in both cases an increase in the parameters yielded a higher accuracy. Yet, the total accuracy difference represented less than 5%. The final values of the parameters were Sigma = 0.01 and C = 100.

3.3. Algorithm evaluation

Table 3 shows the omission, commission and relative bias computed based on the remaining 20% of the training dataset that was used to carry out a cross-tabulation:

The best accuracy was obtained by the C5.0 algorithm followed by SVM and NNs. RF showed the highest commission error and the lowest
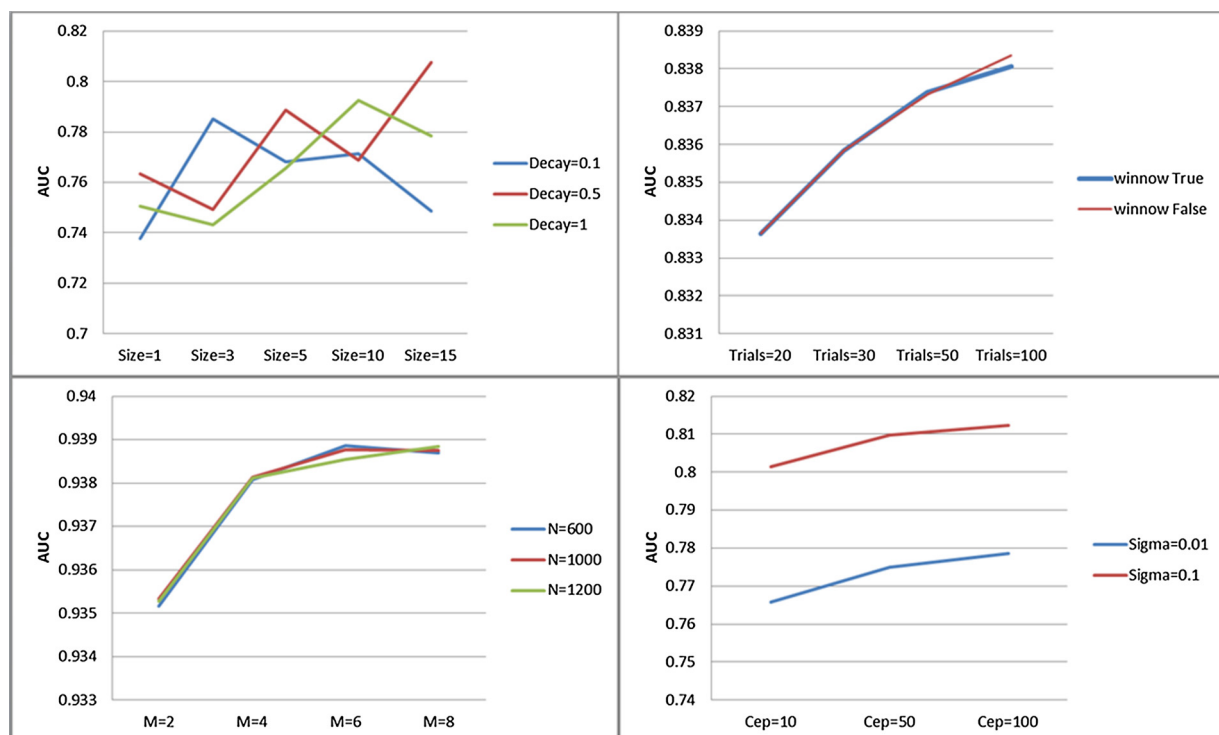

Fig. 5. Variation in AUC in the parameter selection between Neural Net (Top-left), C5 (Top-right), RF (Down-left) and SVM (Down-right).

Table 3
Validation results.

Algorithm	Commission	Omission	Rel. Bias
Random Forest	0.56	0.11	1.01
SVM	0.17	0.36	−0.23
C5.0	0.18	0.32	−0.17
Neural Net	0.21	0.39	−0.23

omission. In other words, whereas SVM, C5 and NN did not detect approximately a 30% of the burned area, RF had many false positives.

3.4. Comparison with existing BA information

The comparison of the outputs of each machine learning algorithm with the official fire database is presented in Table 4. For RF and C5.0 the best results extracted from the sensibility analysis of the probability of burned are shown.

Taking into account the three areas globally, the best results were obtained by C5. This model showed a good tradeoff between commission and omission, with better error rates than MCD45 and comparable to the MCD64 product. In the same line, RF also reached comparable results with C5.0 and MCD64. Despite RF presented higher commission, it showed a better balance between commission and omission errors. NN and SVM showed the worse performance with errors reaching 90% in the case of SVM.

The results of the sensitivity analysis of the probability of burned pixels on RF and C5.0 is also presented in Fig. 6.

For RF, the probability of burned threshold that minimizes the error rates is 40%. This threshold is the same for the three regions, except for California where 30% obtained slightly better results. Nevertheless, the 40% threshold still yielded better results than the MCD45 product. On the other hand, the sensitivity analysis performed for C5.0 did not reveal a clear trend showing a large variation over the range of probability values used.

Table 5 shows the results obtained by comparing the outputs of our algorithms as well as the MDC45 to the MDC64 per region, globally and for Africa.

The results reveal that the MCD45 obtained the closest agreement with the MCD64 product, followed by RF, whereas very poor results were achieved by NN and SVM. Among the analyzed models, RF shows the best balance between errors. Similar trends were observed focusing

on the African test sites. SVM and NN presented very high omission errors reaching hundred percent in some areas like Sudan. Regarding the sensitivity analysis of the burned probability threshold, results are presented in Fig. 7. The analysis reveals that RF presents a stable a cutoff probability around 30–40% for all sites. On the other hand the cut-off probability threshold for C5.0 varies widely, obtaining better results for lower probabilities.

4. Discussion

4.1. Feature selection

Feature selection is an important part of data mining methodology reducing the dimensionality of the data and removing the possible noise of some of the attributes. Merging the results of the three FS methods used, eight attributes were finally selected. Two of them, the hot spot distance and the NIR difference, are the most important ones followed by two widely used burned area indices like NBR and MIRBI. Similarly, the NDWI5 also showed to be important for BA discrimination. The indices selected showed that those based on SWIR spectral region provides the best discrimination capability for classifying burned areas, which agrees with previous studies (Giglio et al., 2013; Roy et al., 2005). The only qualitative factor common to the three filter is the GFED regions, which facilitates the regional adaptation of the algorithm to the different burned conditions.

The RF filter gave more importance to qualitative factors (land cover, GFED regions, and biomes) and it did not discard any of them. This algorithm also gave more importance to factors indirectly related to fires, like the elevation, slope, and aspect. This information is useful to discard fires where the probability of burned is very low, like non-combustible areas or very high places. This is because of how RF algorithm makes the feature selection evaluating at the same time a group of attributes, making the process biased to those features that are indirectly related with burned (Strobl et al., 2007).

The entropy-based filter gave more importance to attributes related to the reflectance, like the individual bands or the vegetation and burned indices. In this case, the filter analyzes the contribution of each attribute individually, and hence none indirectly related variables, like elevation, are considered important. On the other hand, qualitative variables like land cover have been selected by this filter because there are some classes where fire events are improbable (lakes, dessert, etc.).

Logistic regression provides a fast method to select attributes in an

Table 4
Relative bias, Omission and commission errors for Canada, California, and Australia. For C5.0 and RF the best results of the sensitivity analysis of the probability of burned were provided.

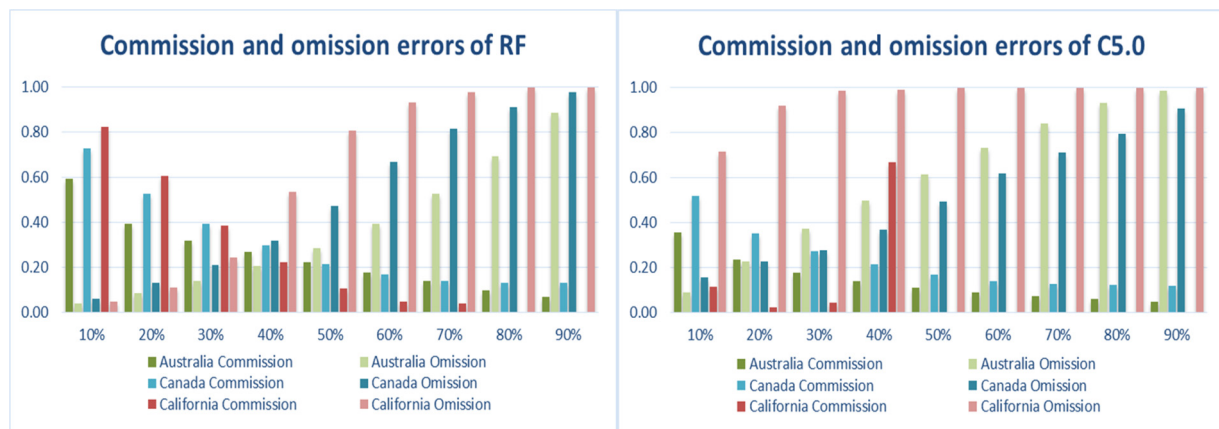| Total	C5.0	RF	NN	SVM	MCD45	MCD64
Commission	0.24	0.27	0.15	0.71	0.11	0.15
Omission	0.24	0.21	0.59	0.90	0.40	0.29
Relative Bias	0.00	0.08	−0.52	−0.64	−0.32	−0.16
California	**C5.0 (prob = 10%)**	**RF (prob = 30%)**	**NN**	**SVM**	**MCD45**	**MCD64**
Commission	0.11	0.38	1	0	0.32	0.13
Omission	0.71	0.24	1	0.98	0.67	0.34
Relative Bias	−0.68	0.22	−0.99	−0.98	−0.52	−0.24
Canada	**C5.0 (prob = 30%)**	**RF (prob = 40%)**	**NN**	**SVM**	**MCD45**	**MCD64**
Commission	0.27	0.30	0.16	0.13	0.21	0.25
Omission	0.28	0.32	0.49	0.99	0.63	0.33
Relative Bias	−0.01	−0.03	−0.40	−0.99	−0.53	−0.11
Australia	**C5.0 (prob = 20%)**	**RF (prob = 40%)**	**NN**	**SVM**	**MCD45**	**MCD64**
Commission	0.24	0.27	0.15	0.72	0.11	0.15
Omission	0.23	0.20	0.59	0.89	0.39	0.29
Relative Bias	0.01	0.09	−0.52	−0.62	−0.31	−0.17

Fig. 6. Sensibility analysis of the probability of burned.

iterative model, where the variables are rejected or included based on the accuracy of a set of models trained using a cross-validation approach. Logistic regression was the most restrictive method, selecting 12 variables from the initial 52, including the GFED regions, four multitemporal variables and the distance matrix of the active fires. This filter showed the same trend as entropy, which selects attributes directly related to fire signal, rejecting those associated with the topography or other environmental factors.

The combination of the three FS methods made the selection of the attributes finally used more objective than ranking methods, for which the cutoff threshold usually is user-defined, thus improving the classification results (Uncu and Türkşen, 2007). Moreover, merging the three methods reduced any potential bias a single method can introduce in the feature selection. This approach is relatively novel in remote sensing and can be especially interesting to apply to high dimensional data such as hyperspectral images.

4.2. Model training

The training dataset was composed of 6.3 million of pixels which

makes it considerably larger than most of the databases used in previous studies in this field (Gómez and Martín, 2011; Özbayoğlu and Bozer, 2012; Pelletier et al., 2016). This resulted in higher computational cost for training the models. In addition, only 0.76% of the instances represented burned areas, i.e. our database was highly unbalanced. This is an important problem because most of the machine learning algorithms assume that the clases are equal distributed, specially when the majority class highly exceeded in number to the minority. Our results showed that RF has a better ability to deal with unbalanced datasets. This is very important for BA mapping globally since BA is a rare event compared to unburned area. Some authors proposed to use disproportionate sampling approach databases to compensate for the shortcomings of random sampling for validating rare events by increasing the proportion of samples collected within the rare class, thus oversampling areas known to be experiencing high rates of fire events via domain knowledge (Farquad and Bose, 2012).

To find the best parameter combination a 10-fold CV approach was used because it results in lower variance than a single hold-out set estimator, which can be very important if the amount of data available is highly unbalanced as in our case (López et al., 2014). The 10-fold CV

Table 5
Relative bias, Omission and commission errors for African test sites. First lines show the error rates performed by the sum of the 6 confusion matrices. For C5.0 and RF the best results of probability analysis were used.

Global	C5	RF	NN	SVM	MCD45
Commission	0.44	0.42	0.05	0.17	0.28
Omission	0.51	0.43	0.75	0.81	0.43
Relative Bias	−0.13	−0.01	1.72	1.70	−0.21

Africa	C5	RF	NN	SVM	MCD45
Commission	0.49	0.43	0.01	0.07	0.11
Omission	0.57	0.40	0.78	0.79	0.40
Relative Bias	−0.16	0.05	2.12	2.12	−0.32

Angola	C5 (prob = 10%)	RF (prob = 30%)	NN	SVM	MCD45
Commission	0.54	0.52	0.06	0.05	0.30
Omission	0.49	0.56	0.96	0.96	0.41
Relative Bias	0.10	−0.08	−0.96	−0.96	−0.15

South Africa	C5 (prob = 20%)	RF (prob = 40%)	NN	SVM	MCD45
Commission	0.41	0.46	0.13	0.03	0.30
Omission	0.45	0.33	0.87	0.94	0.58
Relative Bias	−0.07	0.23	−0.85	−0.94	−0.40

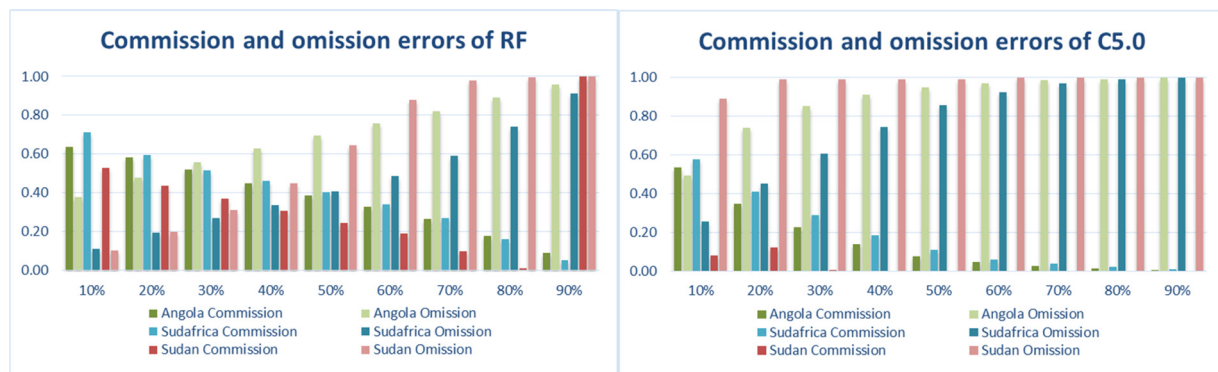| Sudan	C5 (prob = 10%)	RF (prob = 30%)	NN	SVM	MCD45
Commission	0.08	0.37	0.00	0.04	0.34
Omission	0.89	0.31	1.00	0.97	0.34
Relative Bias	−0.88	0.10	−1.00	−0.97	0.01

Fig. 7. Sensibility analysis of the probability of burned in Africa.

validation reduces this variance by averaging over ten different partitions, making the performance estimation less sensitive to the splitting of the data, allowing having a reasonable idea of the model accuracy while avoiding overfitting.

In this study, we trained 41 different models using a 10-CV (410 models in total) to test the sensitivity of each classification algorithm to their parameters. In general, all of them showed low variation (1–6%) with respect to the AUC measure, indicating that for BA classification problems the algorithms were robust to the variation of the parameters. RF showed the least sensitivity to the parameters with a variation in the AUC less than 1%. RF reached almost the same results using only an M value equals to two, i.e. the data partition for classifying burned areas only needed two variables for each tree. This means that the simplest model can obtain comparable results to more complex models. The same trend is observed with the C5.0 algorithm, this effect is related to how boosting decision trees partition the space, making them low sensitive to parameter variations.

The AUC variations due to the change in NN parameters is relatively small (6%), therefore it is desirable to obtain models without a high number of connections among neurons and medium values of decay to avoid overfitting (De Villiers and Barnard, 1993; Karsoliya, 2012). SVM showed little influence of the parameters in the results but it was greatly affected by the size of the training database, resulting in the most computationally expensive of the algorithms. One of the most important issues in training an SVM model is the computation of the distance between all training points, which makes the computational cost very high in comparison with the rest of the algorithms. Thus, whereas the training of RF and C5 took less than 5 h, it took 18 h for the NN and 5 days for the SVM. Training time of NN and particularly SVM could be significantly reduced by performing an instance selection in addition to the FS (Liu and Motoda, 2013).

4.3. Model evaluation

The first evaluation of the models over the remaining 20% of the training dataset revealed very promising results, with omission and commission errors less than 20% and 35%, respectively. Nevertheless, these error rates increased significantly when the results of the classification were compared with the reference datasets. Thus, the SVM model showed very low agreement with the official fire databases or with MCD64. The best results were obtained in Australia where the omission and commission errors reached 89% and 72%, respectively. For the remaining five areas the results show that the model had omission values close to hundred percent. The low generalization ability of the SVM can be related to the sample size used for training (Chen and Lin, 2006). A significant reduction of the database driven by an instance selection can aid to enhance the accuracy of the model (Waske et al., 2010) reducing at the same time the computational cost. The performance of SVM is also affected by other factors like the balance of the training dataset (Eitrich and Lang, 2006; Farquad and Bose,

2012). An unbalanced dataset results in a hyperplane that tends to classify the majority class which is less important than the burned class. To solve this problem instance selection methods like undersampling or oversampling (Farquad and Bose, 2012) could reduce the unbalance of the database improving the estimation of the minority class. Reducing the size of the database by applying instance selection would also make more feasible the estimation of the probability of burned (Tao et al., 2005) instead of the use of a hard classification, which can help to reduce the high omission error (Ramo and Chuvieco, 2017) observed in SVM.

The results obtained by the NN model are slightly better than SVM, improving the error rates in Australia and Canada. For the rest of the areas, the error rates were very high, with omission errors close to 100%. The lack of generalization of the model can be explained by the unbalance of the dataset which resulted in a complex model with a very high number of units in the hidden layers. The unbalance of the training dataset has the same effect as for SVM in the classification, favoring the majority class (unburned), thus increasing the omission errors. Despite two strategies were used to avoid overfitting of the NN, namely the 10-fold cross-validation and the search of optimum parameters, they were not enough to avoid the problems of using an unbalanced database. Recent developments in the field of NNs, particularly the use of newer algorithms like convolutional NN, which is more robust to overfitting (Cheng et al., 2016), could improve our results.

The algorithms based on decision trees showed a different trend to NN and SVM. These algorithms present the following advantages: the option to estimate the probability of burned, the easy parameterization of these models, the possibility to extract the rules that compose the trees, the ability to cope with the unbalanced, and finally, they can deal with big training database. Thus, the C5.0 algorithm showed a good agreement with the official fire database, providing error rates comparable with the MCD64 and MCD45 products. For the test sites located in Africa, the results were also acceptable, obtaining a better balance between omission and commission errors than the MCD45 product, although the commission error was slightly higher. The main problem in the application of this model at global scale is the variability of the probability threshold of burned areas. Figs. 6 and 7 illustrate that the cutoff probability varied between 10–30%, but it did not provide a clear trend to apply at global. A different threshold should be used depending on the biome or region to reduce the error rates. The C5.0 algorithm showed better performance using an unbalanced dataset than SVM and NN but furthermore, it is also possible to apply a cost matrix to take into account the misclassification of the minority class, making it more robust to the unbalance of the training database (Ling and Sheng, 2008).

RF showed the same trend as C5.0 although with some advantages. The probability of burned distribution has a clear convergence point around the 30–40% (see Figs. 6 and 7), therefore it is not necessary a regional adaptation of the threshold. Setting the threshold to 40% the global commission error decreased only by 2% (0.40) and the omission increased by 2% (0.45). This small variation in the accuracy implies

that only areas with low fire occurrence are affected by the threshold.

RF has its own mechanisms to deal with unbalanced datasets, such as the stratification of the training in the tree process creation setting the amount of burned and unburned pixels used to induce the tree. This stratification produces a better representation of the minority class in the forest. Furthermore, it is possible to train random forest using weights to avoid the misclassifications of the minority class. This flexibility of RF along with the easy interpretation of their parameters makes this algorithm more suited to classify burned area at global scale.

Despite of our efforts to prevent some non-desirable effects such as the lack of generalization, some difficulties still remained. One of the most important factors is the training database. Regardless of the great amount of data and the wide spectrum of burned conditions sampled, the impact of unbalanced datasets on the SVM and NN algorithms limited their performance. Procedures such as oversampling in the minority class can help to reduce the bias towards the majority class (C5.0, SVM, NN); nevertheless, each algorithm have their own requirements that have to be considered. Thus, SVM works better if the amount of training data is not big and balanced (Garcia et al., 2011; García et al., 2017), in this case undersampling is also a good choice whereas NN performs better if the number of instances in the training database is high (Maggiori et al., 2017). On the other hand, tree based algorithms are less affected by this issue and yield error rates comparable to those of the MCD45 and MCD64 products. Further efforts need to be done to explore how the training database affects to the performance of each algorithm.

5. Conclusions

This paper evaluated the ability of RF, C5.0, SVM and NN to map burned area in different ecosystems, continents and fire regime conditions, following a data mining approach, which included attribute and parameter selection. Attribute selection was based on merging three different methods, which ensured the importance of the attributes selected avoiding any possible bias of the method used. The most important attributes were related to the reflectance of burned surfaces like the temporal difference of burned area or vegetation indices (MIRBI, NBR, GEMI), given the contrast between burned and unburned areas. Another important factor is related to the detection of active fires that acts as a core to reduce the range of search of the burned pixels. The GFED regions, which are related to the fire behaviour, were also important to map burned areas globally since it enabled accounting for regional differences in fire behavior.

A critical aspect in the performance of the algorithms was the unbalance of the training database since burned area is a rare event representing less than 1% of the data. The application of standard procedures to avoid overfitting such as k-fold cross validation could not avoid the overfitting of all the machine learning methods used. NN and SVM were more affected than C5.0 or RF by the unbalanced database used resulting in complex structures and models that had low generalization ability. Tree-based algorithms showed better performance given their mechanism to deal with large and unbalanced databases as well as their ability to provide the probability of burned. The probabilities extracted from RF were more consistent across regions allowing using a single threshold at global scale. Among the algorithms tested, RF offered the best performance yielding comparable estimates to the MCD64 product in some regions, while for others offered higher commission error rate but with better error balance.

In this context, the development of data mining and machine learning methodologies are very challenging to improve the traditional approaches used for burned area detection. Further research will evaluate the influence of the training dataset over the different machine learning algorithms to compensate the effect of unbalanced training data and overfitting issues.

References

Aldersley, A., Murray, S.J., Cornell, S.E., 2011. Global and regional analysis of climate and human drivers of wildfire. Sci. Total Environ. 409, 3472–3481.

Alonso-Canas, I., Chuvieco, E., 2015. Global burned area mapping from ENVISAT-MERIS and MODIS active fire data. Remote Sens. Environ. 163, 140–152.

Al-Rawi, K., Casanova, J., Calle, A., 2001. Burned area mapping system and fire detection system, based on neural networks and NOAA-AVHRR imagery. Int. J. Remote Sens. 22, 2015–2032.

Archibald, S., Roy, D.P., Van Wilgen, B.W., Scholes, R.J., 2009. What limits fire? An examination of drivers of burnt area in Southern Africa. Glob. Change Biol. 15, 613–630.

Barnes, C.F., Fritz, H., Yoo, J., 2007. Hurricane disaster assessments with image-driven data mining in high-resolution satellite imagery. Ieee Trans. Geosci. Remote Sens. 45, 1631–1640.

Bastarrika, A., Alvarado, M., Artano, K., Martinez, M.P., Mesanza, A., Torre, L., Ramo, R., Chuvieco, E., 2014. BAMS: a tool for supervised burned area mapping using landsat data. Remote Sens. 6, 12360–12380.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31.

Bengio, Y., Grandvalet, Y., 2004. No unbiased estimator of the variance of k-fold cross-validation. J. Mach. Learn. Res. 5, 1089–1105.

Boulila, W., Farah, I.R., Ettabaa, K.S., Solaiman, B., Ghézala, H.B., 2011. A data mining based approach to predict spatiotemporal changes in satellite images. Int. J. Appl. Earth Obs. Geoinf. 13, 386–395.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brewer, C.K., Winne, J.C., Redmond, R.L., Opitz, D.W., Mangrich, M.V., 2005. Classifying and mapping wildfire severity. Photogramm. Eng. Remote Sens. 71, 1311–1320.

Burton, P.J., Parisien, M.-A., Hicke, J.A., Hall, R.J., Freeburn, J.T., 2009. Large fires as agents of ecological diversity in the North American boreal forest. Int. J. Wildland Fire 17, 754–767.

Chen, Y.-W., Lin, C.-J., 2006. Combining SVMs with various feature selection strategies. Feature Extr. 315–324.

Cheng, T., Wang, J., 2008. Integrated spatio-temporal data mining for forest fire prediction. Trans. GIS 12, 591–611.

Cheng, G., Zhou, P., Han, J., 2016. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. IEEE Trans. Geosci. Remote Sens. 54, 7405–7415.

Chu, T., Guo, X., 2013. Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: a review. Remote Sens. 6, 470–520.

Chuvieco, E., Martín, M.P., Palacios, A., 2002. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. Int. J. Remote Sens. 23, 5103–5110.

Chuvieco, E., Riaño, D., Danson, F., Martin, P., 2006. Use of a radiative transfer model to simulate the postfire spectral response to burn severity. J. Geophys. Res. Biogeosci. 111.

Chuvieco, E., Yue, C., Heil, A., Mouillot, F., Alonso-Canas, I., Padilla, M., Pereira, J.M., Oom, D., Tansey, K., 2016. A new global burned area product for climate assessment of fire impacts. Glob. Ecol. Biogeogr. 25, 619–629.

Danielson, J.J., Gesch, D.B., 2011. Global multi-resolution terrain elevation data 2010 (GMTED2010). US Geological Surv.

De Villiers, J., Barnard, E., 1993. Backpropagation neural nets with one and two hidden layers. IEEE Trans. Neural Netw. 4, 136–141.

DeFries, R., Chan, J.C.-W., 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sens. Environ. 74, 503–515.

Deng, S., Katoh, M., Yu, X., Hyyppä, J., Gao, T., 2016. Comparison of tree species classifications at the individual tree level by combining ALS data and RGB images using different algorithms. Remote Sens. 8, 1034.

Dragozi, E., Gitas, I.Z., Stavrakoudis, D.G., Theocharis, J.B., 2014. Burned area mapping using support vector machines and the FuzCoC feature selection method on VHR IKONOS imagery. Remote Sens. 6, 12005–12036.

Eitrich, T., Lang, B., 2006. Efficient optimization of support vector machine learning parameters for unbalanced datasets. J. Comput. Appl. Math. 196, 425–436.

Farquad, M., Bose, I., 2012. Preprocessing unbalanced data using support vector machine. Decis. Support Syst. 53, 226–233.

Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R., 1996. Advances in Knowledge Discovery and Data Mining. AAAI press, Menlo Park.

Gao, B.-C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266.

Garcia, M.L., Caselles, V., 1991. Mapping burns and natural reforestation using thematic mapper data. Geocarto Int. 6, 31–37.

Garcia, M., Chuvieco, E., 2004. Assessment of the potential of SAC-C/MMRS imagery for mapping burned areas in Spain. Remote Sens. Environ. 92, 414–423.

Garcia, M., Riano, D., Chuvieco, E., Salas, J., Danson, F.M., 2011. Multispectral and

LiDAR data fusion for fuel type mapping using support vector machine and decision rules. Remote Sens. Environ. 115, 1369–1379.

García, M., Saatchi, S., Casas, A., Koltunov, A., Ustin, S.L., Ramirez, C., Balzter, H., 2017. Extrapolating Forest canopy fuel properties in the California rim fire by combining airborne LiDAR and landsat OLI data. Remote Sens. 9, 394.

Giglio, L., Randerson, J.T., Werf, G.R., 2013. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). J. Geophys. Res. Biogeosci. 118, 317–328.

Gitelson, A.A., Stark, R., Grits, U., Rundquist, D., Kaufman, Y., Derry, D., 2002. Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction. Int. J. Remote Sens. 23, 2537–2562.

Global Climate Observing System (GCOS), 2004. Implementation Plan for The Global Observing System for Climate in Support of the UNFCC. World Meteorological Organization, Geneva.

Goldammer, J.G., Statheropoulos, M., Andreae, M.O., 2008. Impacts of vegetation fire emissions on the environment, human health, and security: a global perspective. Dev. Environ. Sci. 8, 3–36.

Gómez, I., Martín, M.P., 2011. Prototyping an artificial neural network for burned area mapping on a regional scale in Mediterranean areas using MODIS images. Int. J. Appl. Earth Observ. Geoi. 13, 741–752.

Goswami, S., Chakraborty, S., Ghosh, S., Chakrabarti, A., Chakraborty, B., 2016. A review on application of data mining techniques to combat natural disasters. Ain Shams Eng. J.

Hollmann, R., Merchant, C., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., De Leeuw, G., Forsberg, R., 2013. The ESA climate change initiative: satellite data records for essential climate variables. Bull. Am. Meteorol. Soc. 94, 1541–1552.

Huete, A., 1988. A soil-adjusted vegetation index. Remote Sens. Environ.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: from pixel-based to object-based approaches. ISPRS J. Photogramm. Remote Sens. 80, 91–106.

Jin, Y., Randerson, J.T., Goetz, S.J., Beck, P.S., Loranty, M.M., Goulden, M.L., 2012. The influence of burn severity on postfire vegetation recovery and albedo change during early succession in North American boreal forests. J. Geophys. Res. Biogeosci. 117.

Jurdao, S., Chuvieco, E., Arevalillo, J.M., 2012. Modelling fire ignition probability from satellite estimates of live fuel moisture content. Fire Ecol. 8, 77–97.

Karsoliya, S., 2012. Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. Int. J. Eng. Trends Technol. 3, 714–717.

Khanna, S., Palacios-Orueta, A., Whiting, M.L., Ustin, S.L., Riaño, D., Litago, J., 2007. Development of angle indexes for soil moisture estimation, dry matter detection and land-cover discrimination. Remote Sens. Environ. 109, 154–165.

Klein, Igor, Gessner, Ursula, Kuenzer, C., 2012. Regional land cover mapping and change detection in Central Asia using MODIS time-series. Appl. Geogr. 35, 219–234.

Kloster, S., Mahowald, N., Randerson, J., Lawrence, P., 2012. The impacts of climate, land use, and demography on fires during the 21st century simulated by CLM-CN. Biogeosciences 9.

Kursa, M.B., Jankowski, A., Rudnicki, W.R., 2010. Boruta–a system for feature selection. Fundam. Inf. 101, 271–285.

Lawrence, R.L., Moran, C.J., 2015. The AmericaView classification methods accuracy comparison project: a rigorous approach for model selection. Remote Sens. Environ. 170, 115–120.

Leeuwen, T., Peters, W., Krol, M., Werf, G., 2013. Dynamic biomass burning emission factors and their impact on atmospheric CO mixing ratios. J. Geophys. Res. Atmos. 118, 6797–6815.

Ling, C., Sheng, V., 2008. Cost-sensitive learning and the class imbalance problem. Springer: Encyclopedia of Machine Learning.

Liu, H., Motoda, H., 2013. Instance Selection and Construction for Data Mining. Springer Science & Business Media.

Loboda, T., O'neal, K., Csiszar, I., 2007. Regionally adaptable dNBR-based algorithm for burned area mapping from MODIS data. Remote Sens. Environ. 109, 429–442.

López, V., Fernández, A., Herrera, F., 2014. On the importance of the validation technique for classification with imbalanced datasets: addressing covariate shift when data is skewed. Inf. Sci. 257, 1–13.

Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P., 2017. Convolutional neural networks for large-scale remote-sensing image classification. IEEE Trans. Geosci. Remote Sens. 55, 645–657.

Makido, Y., Shortridge, A., Messina, J.P., 2007. Assessing alternatives for modeling the spatial distribution of multiple land-cover classes at sub-pixel scales. Photogramm. Eng. Remote Sens. 73, 935–943.

Mallinis, G., Koutsias, N., 2012. Comparing ten classification methods for burned area mapping in a Mediterranean environment using landsat TM satellite data. Int.J. Remote Sens. 33, 4408–4433.

Martinez-Fernandez, J., Chuvieco, E., Koutsias, N., 2013. Modelling long-term fire oc-currence factors in Spain by accounting for local variations with geographically weighted regression. Nat. Hazards Earth Syst. Sci. 13, 311–327.

Mitchell, T.M., 1997. Machine Learning. McGraw-Hill, Inc.

Mitri, G.H., Gitas, I.Z., 2013. Mapping post-fire forest regeneration and vegetation re-covery using a combination of very high spatial resolution and hyperspectral satellite imagery. Int. J. Appl. Earth Obs. Geoinf. 20, 60–66.

Moreira, F., Vaz, P., Catry, F., Silva, J.S., 2009. Regional variations in wildfire suscept-ibility of land-cover types in Portugal: implications for landscape management to minimize fire hazard. Int. J. Wildland Fire 18, 563–574.

Mouillot, F., Schultz, M.G., Yue, C., Cadule, P., Tansey, K., Ciais, P., Chuvieco, E., 2014.

Ten years of global burned area products from spaceborne remote sensing—A review: analysis of user needs and recommendations for future developments. Int. J. Appl. Earth Obs. Geoinf. 26, 64–79.

Naidoo, L., Cho, M., Mathieu, R., Asner, G., 2012. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a random Forest data mining environment. ISPRS J. Photogramm. Remote Sens. 69, 167–179.

Nieto, H., Aguado, I., Garcia, M., Chuvieco, E., 2012. Lightning-caused fires in Central Spain: development of a probability model of occurrence for two Spanish regions. Agric. For. Meteorol. 162, 35–43.

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. Manage. 275, 117–129.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., 2001. Terrestrial ecoregions of the world: a new map of life on earth a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience 51, 933–938.

Özbayoğlu, A.M., Bozer, R., 2012. Estimation of the burned area in forest fires using computational intelligence techniques. Procedia Comput. Sci. 12, 282–287.

Padilla, M., Stehman, S.V., Chuvieco, E., 2014. Validation of the 2008 MODIS-MCD45 global burned area product using stratified random sampling. Remote Sens. Environ. 144, 187–196.

Padilla, M., Stehman, S.V., Ramo, R., Corti, D., Hantson, S., Oliva, P., Alonso-Canas, I., Bradley, A.V., Tansey, K., Mota, B., 2015. Comparing the accuracies of remote sen-sing global burned area products using stratified random sampling and estimation. Remote Sens. Environ. 160, 114–121.

Palacios-Orueta, A., Khanna, S., Litago, J., Whiting, M.L., Ustin, S.L., 2006. Assessment of NDVI and NDWI spectral indices using MODIS time series analysis and development of a new spectral index based on MODIS shortwave infrared bands. In: Proceedings of the 1st International Conference of Remote Sensing and Geoinformation Processing. Trier, Germany. pp. 207–209. http://ubt.opus.hbz-nrw.de/volltexte/2006/362/pdf/03-rgldd-session2.pdf.

Parisien, M.-A., Peters, V.S., Wang, Y., Little, J.M., Bosch, E.M., Stocks, B.J., 2006. Spatial patterns of forest fires in Canada, 1980–1999. Int. J. Wildland Fire 15, 361–374.

Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. Assessing the ro-bustness of random forests to map land cover with high resolution satellite image time series over large areas. Remote Sens. Environ. 187, 156–168.

Pereira, A.A., Pereira, J., Libonati, R., Oom, D., Setzer, A.W., Morelli, F., Machado-Silva, F., de Carvalho, L.M.T., 2017. Burned area mapping in the brazilian savanna using a one-class support vector machine trained by active fires. Remote Sens. 9, 1161.

Petropoulos, G.P., Vadrevu, K.P., Xanthopoulos, G., Karantounias, G., Scholze, M., 2010. A comparison of spectral angle mapper and artificial neural network classifiers combined with landsat TM imagery analysis for obtaining burnt area mapping. Sensors 10, 1967–1985.

Picton, P., 2000. Introduction to Neural Networks, second ed. Macmillan Publishers Limited.

Pinty, B., Verstraete, M., 1992. GEMI: a non-linear index to monitor global vegetation from satellites. Vegetatio 101, 15–20.

Plummer, S., Arino, O., Ranera, F., Tansey, K., Chen, J., Dedieu, G., Eva, H., Piccolini, I., Leigh, R., Borstlap, G., 2005. The GLOBCARBON initiative: multi-sensor estimation of global biophysical products for global terrestrial carbon studies. Envisat ERS Symposium.

Quinlan, J., 1993. C 4.5: Programs for Machine Learning. The Morgan Kaufmann Series in Machine Learning. Morgan Kaufmann, San Mateo, CA 1993.

Quintano, C., Fernández-Manso, A., Stein, A., Bijker, W., 2011. Estimation of area burned by forest fires in Mediterranean countries: a remote sensing data mining perspective. For. Ecol. Manage. 262, 1597–1607.

Ramo, R., Chuvieco, E., 2017. Developing a random forest algorithm for MODIS global burned area classification. Remote Sens. 9, 1193.

Rogan, J., Franklin, J., 2001. Mapping wildfire burn severity in southern California for-ests and shrublands using enhanced thematic mapper imagery. Geocarto Int. 16, 91–106.

Roy, D., Jin, Y., Lewis, P., Justice, C., 2005. Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. Remote Sens. Environ. 97, 137–162.

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.-P., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83, 135–148.

Schneider, P., Roberts, D., Kyriakidis, P., 2008. A VARI-based relative greenness from MODIS data for computing the fire potential index. Remote Sens. Environ. 112, 1151–1167.

Schoennagel, T., Nelson, C.R., Theobald, D.M., Carnwath, G.C., Chapman, T.B., 2009. Implementation of national fire plan treatments near the wildland–urban interface in the western United States. Proc. Natl. Acad. Sci. 106, 10706–10711.

Smith, A., Drake, N., Wooster, M., Hudak, A., Holden, Z., Gibbons, C., 2007. Production of landsat etm + + reference imagery of burned areas within Southern African sa-vannahs: comparison of methods and application to MODIS. Int. J. Remote Sens. 28, 2753–2775.

Strobl, C., Boulesteix, A.-L., Augustin, T., 2007. Unbiased split selection for classification trees based on the Gini index. Comput. Stat. Data Anal. 52, 483–501.

Stroppiana, D., Grégoire, J.M., Pereira, J.M.C., 2003. The use of SPOT VEGETATION data in a classification tree approach for burnt area mapping in Australian savanna. Int. J. Remote Sens. 24, 2131–2151.

Tansey, K., Grégoire, J.M., Defourny, P., Leigh, R., Pekel, J.F., Van Bogaert, E., Bartholomé, E., 2008. A new, global, multi-annual (2000–2007) burnt area product

at 1 km resolution. Geophys. Res. Lett. 35.

Tao, Q., Wu, G.-W., Wang, F.-Y., Wang, J., 2005. Posterior probability support vector machines for unbalanced data. IEEE Trans. Neural Netw. 16, 1561–1573.

Traore, B.B., Kamsu-Foguem, B., Tangara, F., 2017. Data mining techniques on satellite images for discovery of risk areas. Expert Syst. Appl. 72, 443–456.

Trigg, S., Flasse, S., 2001. An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah. Int. J. Remote Sens. 22, 2641–2647.

Uncu, Ö., Türkşen, I., 2007. A novel feature selection approach: combining feature wrappers and filters. Inf. Sci. 177, 449–466.

van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G., Mu, M., Kasibhatla, P.S., Morton, D.C., DeFries, R., Jin, Yv., van Leeuwen, T.T., 2010. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). Atmos. Chem. Phys. 10, 11707–11735.

Vapnik, V., 2013. The Nature of Statistical Learning Theory. Springer science & business media.

Waske, B., Benediktsson, J.A., 2007. Fusion of support vector machines for classification of multisensor data. IEEE Trans. Geosci. Remote Sens. 45, 3858–3866.

Waske, B., van der Linden, S., Benediktsson, J.A., Rabe, A., Hostert, P., 2010. Sensitivity of support vector machines to random feature selection in classification of hyperspectral data. IEEE Trans. Geosci. Remote Sens. 48, 2880–2889.

Wittenberg, L., Malkinson, D., Beeri, O., Halutzy, A., Tesler, N., 2007. Spatial and temporal patterns of vegetation recovery following sequences of forest fires in a Mediterranean landscape, Mt. Carmel Israel. Catena 71, 76–83.

Yang, X., Liu, Z., 2005. Using satellite imagery and GIS for land-use and land-cover change mapping in an estuarine watershed. Int. J. Remote Sens. 26, 5275–5296.

Zhou, F., Zhang, A., Townley-Smith, L., 2013. A data mining approach for evaluation of optimal time-series of MODIS data for land cover mapping at a regional level. ISPRS J. Photogramm. Remote Sens. 84, 114–129.