

Article

Developing a Random Forest Algorithm for MODIS Global Burned Area Classification

Rubén Ramo and Emilio Chuvieco *

Enviromental Remote Sensing Research Group, Department of Geology, Geography and the Environment, University of Alcalá, Colegios 2, 28801 Alcalá de Henares, Spain; ruben.ramo@uah.es

* Correspondence: emilio.chuvieco@uah.es; Tel.: +34-918854438

Received: 7 October 2017; Accepted: 16 November 2017; Published: 21 November 2017

Abstract: This paper aims to develop a global burned area (BA) algorithm for MODIS BRDF-corrected images based on the Random Forest (RF) classifier. Two RF models were generated, including: (1) all MODIS reflective bands; and (2) only the red (R) and near infrared (NIR) bands. Active fire information, vegetation indices and auxiliary variables were taken into account as well. Both RF models were trained using a statistically designed sample of 130 reference sites, which took into account the global diversity of fire conditions. For each site, fire perimeters were obtained from multitemporal pairs of Landsat TM/ETM+ images acquired in 2008. Those fire perimeters were used to extract burned and unburned areas to train the RF models. Using the standard MD43A4 resolution (500×500 m), the training dataset included 48,365 burned pixels and 6,293,205 unburned pixels. Different combinations of number of trees and number of parameters were tested. The final RF models included 600 trees and 5 attributes. The RF full model (considering all bands) provided a balanced accuracy of 0.94, while the RF RNIR model had 0.93. As a first assessment of these RF models, they were used to classify daily MCD43A4 images in three test sites for three consecutive years (2006–2008). The selected sites included different ecosystems: Australia (Tropical), Boreal (Canada) and Temperate (California), and extended coverage (totaling more than $2,500,000$ km 2). Results from both RF models for those sites were compared with national fire perimeters, as well as with two existing BA MODIS products; the MCD45 and MCD64. Considering all three years and three sites, commission error for the RF Full model was 0.16, with an omission error of 0.23. For the RF RNIR model, these errors were 0.19 and 0.21, respectively. The existing MODIS BA products had lower commission errors, but higher omission errors (0.09 and 0.33 for the MCD45 and 0.10 and 0.29 for the MCD64) than those obtained with the RF models, and therefore they showed less balanced accuracies. The RF models developed here should be applicable to other biomes and years, as they were trained with a global set of reference BA sites.

Keywords: burned area; Random Forest; MODIS

1. Introduction

Biomass burning is a relevant factor of the Earth system, affecting vegetation productivity, land use, and atmospheric emissions [1–3]. It has social implications as well, impacting people's lives and properties, particularly in developed countries where urban areas are intermixed with forests [4].

The mutual influences of fire and climate explain that Fire Disturbance is considered one of the Essential Climate Variables (ECVs) by the Global Climate Observing System (GCOS) [5]. Accordingly, several space agencies are aiming to develop systematic assessments of fire occurrence and fire impacts as part of the goal of improving the use of satellite data in climate modelling. This is the main purpose of the Climate Change Initiative (CCI) program of the European Space Agency (ESA) [6]. Within this program, the Fire_cci project aims to develop long-term time series of burned areas

(<http://www.esa-fire-cci.org/>, last accessed 15 February 2017), adapted to the needs of climate modelers [7,8].

Burned area (BA) mapping from satellite images has been undertaken in recent decades using a wide set of methods and sensors, including coarse-resolution data (Advanced very high resolution radiometer (AVHRR), Moderate-Resolution Imaging Spectroradiometer (MODIS), or VEGETATION [9–11]), data at medium resolution (Landsat TM/ETM or SPOT-HRV [12–14]), and very high-resolution data (Worldview, Ikonos: [15,16]).

A few of these BA products provide global coverage. They are particularly useful as input to dynamic global vegetation models [7] and gas emission estimations [17]. The most used products for these modeling efforts were derived from the MODIS sensor on board the Terra and Aqua satellites. The standard MODIS BA product was named MCD45, combining both satellites to estimate BA from a predictive reflectance model [11]. An alternative BA product, named MCD64, was originally derived for the Global Fire Emissions Database [18], but now it has become the standard MODIS BA product in collection 6, released in 2017. It relies on a hybrid algorithm combining active fire and reflectance changes [19]. Both MODIS BA products use the 500 × 500 m MODIS bands, with coverage ranging from visible to the short-wave infrared (SWIR).

From the ESA side, different BA products have also been produced in the last decade: the L3JRC [10] and Globcarbon [20], based on SPOT-VEGETATION data at 1000 × 1000 m spatial resolution, and the Fire_cci product based on ENVISAR MERIS images [8,21] at 300 × 300 m.

BA algorithms are very diverse. Most of the time, they aim to discriminate BA in local-regional conditions, and therefore are difficult to generalize to other sites with different fire characteristics. Considering the wide variety of burning conditions worldwide, it is very challenging to design a global BA algorithm. In some cases, these algorithms rely on physical principles that are adapted locally [19,21]. In other cases, such as, for instance, the L3JRC product [10], several regional algorithms are used in different ecosystems and the final product is a merging of regional outputs, which may have quite different accuracies [22].

Random Forest [23] is one of the most frequently used algorithms for classification of satellite images [24,25]. Its potential to integrate data from different scales and sources, and its robustness to noise and non-normal distributions explains the wide use of RF for many satellite image applications. For instance, RF has been successfully used in land-cover classification [26–28], discrimination of forest species [29,30], agricultural crop classification [31], classification of hyperspectral data [32], biomass estimation and quantifying forest structures [33,34]. The algorithm has also been tested to integrate information acquired by different sensors [35]. RF has also been applied for forest fire research, particularly for fire occurrence prediction [36], as well as to determine factors of fire severity [37] and characterize fire regimes [38,39], but it has not yet been tested for burned area discrimination.

Based on the success of RF models for land-cover mapping and the robustness of this approach to coping with a wide variety of input attributes, the main objective of this paper was to design a BA classification based on RF models. We trained them from a global sample of burned areas derived from Landsat images. Two RF models were generated: a full model, considering all 7 MODIS optical bands (RF Full model), as well as spectral indices and auxiliary variables; and another one restricting the MODIS bands to the NIR and R (RF RNIR model). The SWIR band, in conjunction with the NIR, is more sensitive to burned areas [40,41]. This second model aimed to test the potential application of RF to detect BA from the 250 m resolution bands of MODIS (bands 1 and 2), as well as from other sensors retrieving information just on the R-NIR space (such as MERIS, OLCI, DMC, etc.). As a first assessment of the performance of the RF models in different fire regimes, they were used to classify BA in three large areas, located in Tropical, Temperate and Boreal regions. We compared BA detections in the three test areas with national fire perimeters, as well as with two existing BA MODIS products.

2. Methods

2.1. Algorithm Structure

The development of the RF models was based on three phases (Figure 1). The first one aimed to develop the training database, extracted from MODIS images using reference BA perimeters derived from Landsat data. The second phase trained the RF models by selecting the most appropriate set of classification parameters; and the third one tested the performance of the models in three different regions. All input images and auxiliary information used in this study were reprojected to WGS84 to guarantee spatial consistency among them. RF models were generated using public domain code (RandomForest R package).

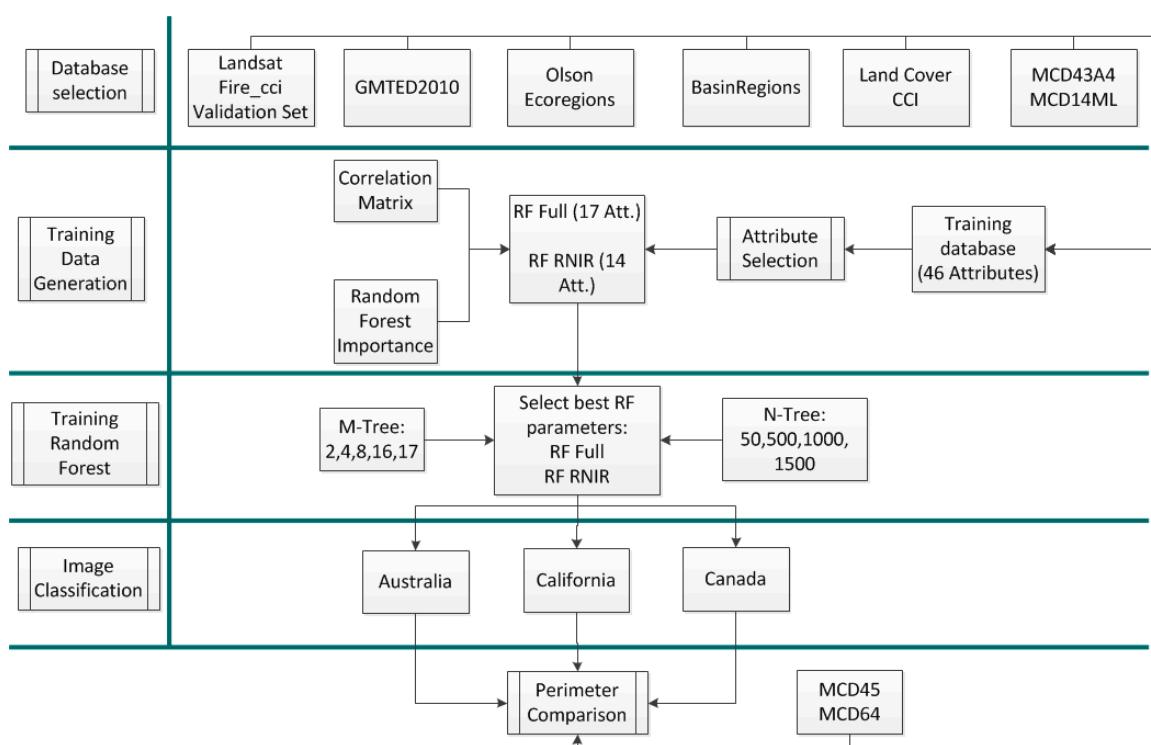


Figure 1. Algorithm flowchart.

2.2. Input Data

Input data for the RF algorithm was the last available version of the MCD43A4 product (v6) (https://lpdaac.usgs.gov/dataset_discovery/MODIS/MODIS_products_table/mcd43a4_v006, last access on August 2017), based on the Schaaf [42] BRDF inversion algorithm. This product combines the observations from Terra and Aqua satellites and daily provides 500 m MODIS BRDF corrected reflectances in bands 1 to 7 (covering the visible to the SWIR spectrum: Table 1). The current version of the product includes daily images, where the corrected reflectance corresponds to the BRDF inversion of a moving 16-day period, including only cloud-free observations. The date of the product is centered upon the retrieval period. We used full inversion and magnitude inversion pixel values (the latter when there are at least seven observations in the moving 16-day window), which correspond to quality values 0, 1, and 2. We followed a conservative approach to avoid losing too many observations in cloudy or poorly observed areas [43].

In addition to the reflectance bands, the following spectral indices were computed as inputs to the training phase.

Soil Adjusted Vegetation Index (SAVI): Designed to improve the sensitivity of the vegetation index to the contribution of soil reflectance [44], and computed as:

$$SAVI = \frac{\rho_{NIR} - \rho_{Red}}{(\rho_{NIR} + \rho_{Red} + L)}(1 + L) \quad (1)$$

where ρ_{NIR} is band 2 and ρ_{Red} is band 1 from MODIS. L is 0.5 [45]. SAVI has been used in burned land mapping by several authors, as it improves the discriminability of BA in sparsely vegetated areas [46–48].

Table 1. MCD43A4 characteristics.

Band	Bandwidth (nm)
1—Red	620–670
2—NIR	841–876
3—Blue	459–479
4—Green	545–479
5—SWIR1	1230–1250
6—SWIR2	1628–1652
7—SWIR3	2105–2155

Global Environmental Monitoring Index (GEMI): This index was developed by Pinty and Verstraete [49] as a more global index of vegetation activity than NDVI. Being a non-linear index, several studies have found it more sensitive to low reflectances in the R/NIR space. For this reason, it has also proven more appropriate than NDVI for burned area detection [47,50]:

$$GEMI = \eta(1 - 0.25\eta) - \frac{(\rho_{Red} - 0.125)}{1 - \rho_{Red}} \quad (2)$$

$$\eta = \frac{2(\rho_{NIR}^2 - \rho_{Red}^2) + 1.5\rho_{NIR} + 0.5\rho_{Red}}{\rho_{Red} + \rho_{NIR} + 0.5} \quad (3)$$

With the same meaning as (1).

Normalized Burn Ratio (NBR): First proposed by Lopez and Caselles [51], and widely used for analyzing burn severity [52,53]:

$$NBR = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (4)$$

For our analysis, ρ_{NIR} and ρ_{SWIR} are MODIS bands 2 and 7.

Normalized difference water index (NDWI): This has the following structure:

$$NDWI = \frac{\rho_{SWIR} - \rho_{NIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (5)$$

It was originally designed for water content estimation [54], and it has been shown to be closely related to fuel moisture content [55], and is therefore associated with short-term fire impacts (reduction of leaf water caused by fire heat [56]). The index variables have the same meaning as NBR, but using the MODIS bands 5 and 6 for the SWIR channels.

Visible Atmospherically Resistant Index (VARI): Designed by Gitelson [57] for agricultural purposes, but has since been used in several fire-related studies, particularly to improve detection of chlorophyll changes as a result of fire [58]:

$$VARI = \frac{\rho_{Green} - \rho_{Red}}{\rho_{Green} + \rho_{Red} - \rho_{Blue}} \quad (6)$$

where ρ_{Green} is band 4, ρ_{Blue} is band 3 and ρ_{Red} is band 1 from MODIS.

Enhanced vegetation index (EVI): this has proven more sensitive than NDVI for monitoring densely vegetated areas and reducing atmospheric effects [45]:

$$EVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + 6 \cdot \rho_{Red} - 7.5 \cdot \rho_{Blue} + 1} \quad (7)$$

where ρ_{NIR} is band 2, ρ_{Blue} is band 3, and ρ_{Red} is band 1 from MODIS. This index was included as a more sensitive and robust estimation of pre-fire vegetation conditions than the standard NDVI, but it has also been successfully tested in post-fire regeneration analysis [59].

Mid-Infrared Burnt Index (MIRBI): Designed by Trigg [41] to estimate burning impacts in South Africa, and later successfully used for BA classification in other ecosystems [60].

$$MIRBI = 10 \cdot \rho_{SWIR3} - 9.8 \cdot S \cdot \rho_{SWIR2} + 2 \quad (8)$$

where ρ_{SWIR2} and ρ_{SWIR3} are reflectances of band 6 and band 7, respectively.

In addition to reflectance and spectral indices, we also used the MODIS thermal anomalies product (MCD14ML version 5.1), which provides daily and global information on hotspots (HS). Most of these are active fires, and therefore are very useful for discriminating burned patches from other land-cover changes. Since the thermal contrast of fires with the background is much higher than the reflectance change caused by the burned signal, HS information has been widely used in BA algorithms [61,62], particularly at a global scale [19,21]. For this research, we created an auxiliary variable with the distances from each pixel to the closest HS. To compute this distance, we also considered HS located in the surrounding margins of the Landsat sampled scenes (up to 50 km outside), to avoid artifacts created by border effects.

We included auxiliary data in the training phase that could help to adapt the RF models to regional BA conditions, and had been related to fire occurrence in previous studies: elevation, biomes and continental regions [63,64]. We imported elevation data from the Global multi-resolution terrain elevation data 2010 (GMTED2010) [65]. Slopes and aspect were calculated using formulas proposed by Burrough [66]. Land cover information was obtained from the Land Cover CCI project, which provides global LC information at 300 m resolution in three different epochs: 2000, 2005, 2010 [67]. This variable was produced from ENVISAT-MERIS images, and follows the standard FAO classification system [68]. We used only the 2005 epoch, as it was the closest pre-fire period to the training and test series used in this research. Ecosystem variation was considered through the Olson biomes map [69], which divides the world into 16 regions on the basis of their geology, climate and evolutionary history. We also incorporated the continental regions defined for the Global Fire Emission Database (GFED), which have previously been used to quantify fire effects, into the input database [18].

2.3. Training Phase

The RF models are built from iterative generations of decision trees, using a random sub-dataset of the input data for each one [23]. Those trees are later used to classify the target area, where each pixel is generally assigned to the category with maximum number of single-tree assignments. The user, depending on data complexity, decides the number of trees. A similar amount of input data is sampled in each random selection. Each of these samples is divided into two groups: a training set, composed by two-thirds of the sampled data; and a validation set (named Out of bag, OOB), which is used to validate the decision tree generated in each step (OOB error). Each decision tree is built with a randomly selected group of attributes (M parameter). To create a decision tree, the algorithm looks for the best data partition [70] using the randomly selected attributes. The quality of the training phase is estimated from the average OOB error [71]. When the training phase is over, every case in the training set is classified from the ensemble of all decision trees previously generated, assigning it to the category with the majority of assignments.

In order to take into account a great variety of burning conditions, the training phase was based on a statistically designed global sample of BA perimeters. This sample was originally generated within the Fire_cci project to validate a global BA product [72]. More specifically, 130 sites were selected using two-stage stratified random sampling. The first stratification level was based on the Olson biomes (seven classes), and the second one on the BA extent in 2008 (high and low), provided by the Global Fire Emissions Database (GFED) version 3 [18,19]. The global distribution of the sampled sites is illustrated in Figure 2. Each site covers approximately the area of a Landsat scene.

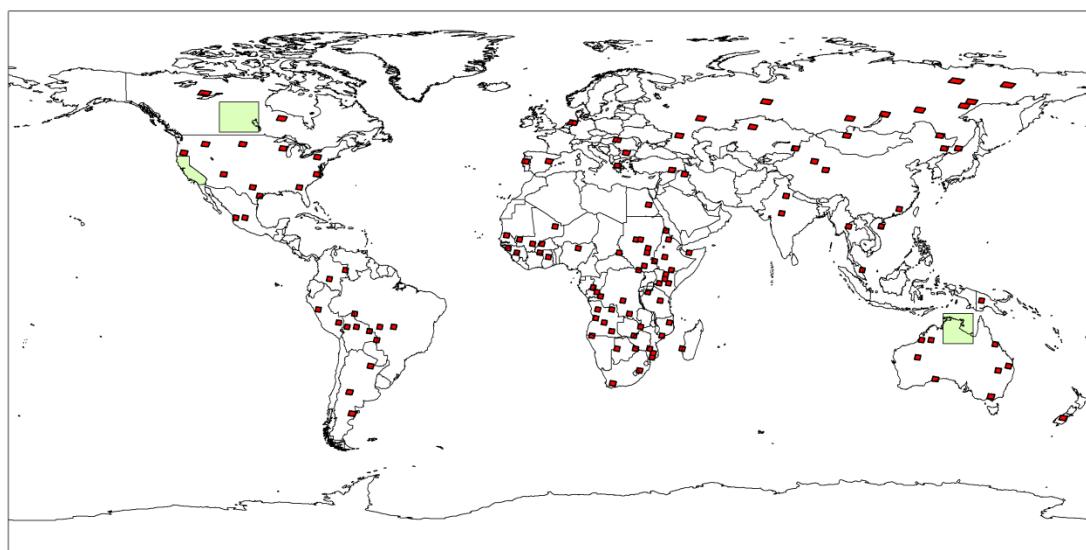


Figure 2. Distribution of training (red) and test areas (green).

Reference burned perimeters were extracted from Landsat multitemporal images using a semi-automatic algorithm [60], and were visually inspected by two independent interpreters to assure data quality [72]. Unburned islands within fire perimeters were discriminated as well. All Landsat images were acquired in 2008. Derived perimeters were restricted to those burned areas occurring in between the two Landsat dates for each scene following standard CEOS Cal-Val approaches. Each burned pixel was dated with the date of the closest HS. The 130 sampled scenes included a total of 66,717 burned patches. Being developed from a statistically designed sample, the training sites can be considered a reliable estimation of fire conditions worldwide, at least for 2008; and therefore, the derived RF models should be globally applicable. The sample includes 1.58 million km^2 , of which 31,578 km^2 were burned. The most extended land covers in the sample were: Rainfed cropland (10.10%); Tree cover, broadleaved, evergreen, closed to open >15% cover proportion (10.63%); Tree cover, broadleaved, deciduous, open 15–40% (5.45%); Tree cover, needleleaved, evergreen, closed to open >15% (8.54%); Shrubland (14.42%); Grassland (16.16%); Sparse vegetation (tree, shrub, herbaceous cover <15%) (6.95%); and Water bodies (5.13%).

The RF training database was created by overlapping a 500×500 m grid size (MODIS pixel resolution) to the Landsat scenes, computing the proportion of burned area in each cell (from 0 to 100%). Pixel size was homogenized to the 500×500 m spatial resolution of the MCD43A4 product. A total of 15,000 images were necessary to generate this training dataset for two reasons: (1) MODIS tiles had to be selected from all regions covered by the Landsat scenes; and (2) all MODIS daily images covering the two Landsat observations had to be processed, as the burning date of each pixel may occur anytime in between those dates. In addition, the temporal series included images acquired from at least 10 days before and 10 days after the Landsat acquisition dates, to avoid problems caused by cloud obscuration or BRDF inversion issues.

For each cell of the training dataset, the MODIS reflectances were obtained based on the most probable date of the burned perimeters. For doing so, we had first to date every burned patch. Since the Landsat images in each site were separated from 16 to 144 days, depending on the image availability (although most scenes had less than 32 days of separation between Landsat pairs [72]), we used the date of the active fires (extracted from the MCD14 product) to label each pixel within the Landsat burned perimeters. Once the date was assigned to each 500×500 m cell, we extracted MODIS reflectances from the day prior to the fire (t1) and two days after the fire (t2). We did not select as t2 the day just after the fire in order to avoid the potential contamination caused by the smoke plumes. Whenever the target dates didn't have good observations, due to either image gaps, noise, or clouds (which are removed in the MCD43A4 product), we selected alternative days within a 10-day temporal window before t1 and after t2. For non-burned pixels, the t1 reflectance was taken from the MODIS image of the first date of the Landsat multitemporal pair. The t2 reflectance was extracted from the MODIS image dated in the median day between the two Landsat images.

To reduce potential noises caused by errors in dating or the impact of previous burnings (those occurring before the first date of the Landsat pair), we introduced three filters in the training dataset. Firstly, we removed pixels labelled as burned when the NIR reflectance of the t2 image was higher than t1, as it is extremely unlikely that a burned pixel exhibits an increase in comparison to the pre-fire NIR reflectance [73]. We tested this issue in many images, and found that all these situations corresponded to false detections. We also removed those pixels within a 3000 m radius of any HS in all images acquired 90 days before t1. This decision was intended to avoid the inclusion of pixels that were burned in previous periods as unburned pixels. Finally, those pixels with less than 80% burned area were removed as well. We applied this criterion in order for training to take place only with those pixels that were more clearly burned, avoiding pixels with a mixed signal.

Applying these rules, our final training database had a total of 6,341,570 pixels (500×500 m spatial resolution), including 48,365 burned pixels (0.76%) and 6,293,205 unburned. This sample size is considerably larger than most previous studies based on RF classification models [24,26,28,38,74].

The attributes of the training database included t1 and t2 reflectances for all 7 bands of the MODIS MCD43 product, plus the spectral indices and their temporal differences (Table 2). In addition to these, seven auxiliary variables were considered (Land Cover, Olson Biomes, GFED Regions, HS distance, Elevation, Slope, and Aspect), totaling 46 attributes for the training dataset.

Table 2. Input attributes for training the RF models.

Attributes	Number
MODIS Bands 1 to 7. Prefire and Postfire	14
SAVI: Prefire, Postfire, Difference	3
GEMI: Prefire, Postfire, Difference	3
NBR: Prefire, Postfire, Difference	3
VARI: Prefire, Postfire, Difference	3
NDWI: Prefire, Postfire, Difference. Using B5, B6	6
EVI: Prefire, Postfire, Difference	3
MIRBI: Prefire, Postfire, Difference	3
NIR temporal difference	1
MODIS active fires: Distance matrix	1
Auxiliary variables: Elevation, slope, aspect, land cover, Olson biomes, GFED regions	6

2.4. Attribute Selection

The RF algorithm assesses the importance of the different input layers in the construction of the decision trees by measuring the variation of the Gini index through the OOB error when an attribute is removed. The Gini coefficient measures error balance. The lower the value the lower the relevance of the parameter as it implies a marginal change in total error [70]. This process was repeated for all trees and attributes to obtain an average importance for each of the attributes in the RF model.

The most relevant variables are those whose absence greatly increases the total error. The absolute values were converted to percentages to standardize the outputs. We also computed the correlation matrix between quantitative input variables. Even though the impact of correlated variables on RF models is not yet fully demonstrated [25,75], we removed highly correlated variables ($r > 0.8$) to avoid redundancy. To select the most explicative attributes, a RF model with a high number of trees ($N = 1500$) was used. Each tree had 7 attributes, which is close to the default: $M = \text{SQRT}(\#\text{attributes})$, as suggested by Breiman [23].

2.5. Parameter Selection

Once the input attributes were selected, the next step was finding the best parameters to build the RF models. Several trainings were conducted after modifying two factors: number of trees and number of input variables [76]. In both cases, the final selection was driven by both the accuracy and performance of the output models. In all cases, the training dataset was divided in two groups: calibration (80% of the data) and test (the remaining 20%). We selected this threshold instead of the standard 2/3 and 1/3 proportion in order to increase the number of burned pixels in the calibration, as they were a small proportion of the training sample. However, we also tested the standard proportions, with very similar accuracy values. Finally, considering the high imbalance of the training dataset, with burned areas only covering less than 1% of all training data, we performed a stratified training [77–79], forcing to introduce in each tree at least 10% of randomly selected burned pixels. Doing this ensured that all burned pixels were included in the RF model [80].

A total of 24 RF models were built. We tested combinations of four numbers of trees ($N = 50, 500, 1000, 1500$) and six numbers of attributes ($M = 2, 4, 8, 12, 16, 17$). The models were assessed by the Balanced Accuracy [81] metric defined in (Table 3):

$$\text{Balanced Accuracy} = \frac{\frac{E_{11}}{(E_{11}+E_{21})} + \frac{E_{22}}{(E_{22}+E_{12})}}{2} \quad (9)$$

Table 3. Structure of the BA confusion matrix.

		Reference	
Predicted	Burned	Non-Burned	Row Total
Burned	E_{11}	E_{12}	E_{1+}
Non-burned	E_{21}	E_{22}	E_{2+}
Col. total	E_{+1}	E_{+2}	E_{Σ}

In terms of performance, the literature suggests using the simplest model (less #trees and #parameters) if they provide similar accuracy to more complex ones [28,81].

2.6. Classification Phase

The final stage of the process was the application of the RF model to the daily MCD43A4 images to classify burned areas in different test sites. Since we had t_1 and t_2 in the RF models, we had to select two daily images for each classification. We identified the first day of each daily series as t_1 , and the third day as t_2 , and moved this temporal window iteratively throughout the full year. For each day and pixel, the RF model computed N classifications (as many as the number of trees in the model), each one built with M attributes. The RF classification output is the number of times the trees classify that pixel as burned or unburned. Since the burned category is quite unbalanced over the unburned (very low proportion of burned pixels), we tested several thresholds to assign each pixel to the burned category. The range tested included from 20% to 80% of the classification trees. We finally selected the threshold that provided the best balance between omission and commission errors [81].

At the end of the process, all BA detections in the daily images were merged by summing up all dates of detection (0 if unburned) to create an annual composite of burned areas. Other compositing periods (monthly, for instance) may be created, particularly for those Tropical regions that may burn more than once within the same year. Finally, a modal filter of 3×3 pixels was applied to improve the spatial coherence of the resulting product. Three years were classified—2006 to 2008—to test the temporal consistency of the RF models.

2.7. Alternative Model Based on Red/NIR Bands

As previously stated, we built a second model with just the MODIS Red and NIR bands, mimicking the spectral properties of the MODIS higher resolution product (250 m), but which was also potentially applicable to other sensors limited to the R/NIR space (OLCI, MERIS, AVHRR, etc.). The RF RNIR model was created from the input reflectances of MODIS bands 1 and 2, the spectral indices derived from those bands, and the same auxiliary variables used for the RF Full model. The RF RNIR model was trained with the same parameters of the RF full model, selecting the most relevant attributes within the restricted list of input variables.

2.8. Comparison with Existing BA Information

As a first assessment of the accuracy of the RF models, the outputs were compared with national fire perimeters and existing global BA MODIS products. Three sites were selected for comparison, as they had available reference perimeters and were representative sites of three different fire regimes (Boreal, Temperate and Tropical).

As an example of Boreal fire conditions, we selected a region of 926,167 km² in Canada, covering the provinces of Manitoba and Saskatchewan. Fire perimeters were downloaded from the Canadian Wildland Fire Information System (cwfis.cfs.nrcan.gc.ca/ha/nfdb/, last accessed August 2017). This system is part of the Canadian National Fire Database (CNFDB), which is a collection of fire data provided by different fire management agencies (provinces, territories, and Parks Canada), and generated by different means. The CNFDB has been extensively used for fire regime characterization studies [82–85]. Fire perimeters were reprojected to WGS84 to facilitate cross-tabulation analysis with our results.

For Tropical fires, we selected a region in Northern Australia. The area embraces a region of 1,192,585 km² mainly covered by Savannah and Tropical forest. Fire perimeters were downloaded from the North Australian Fire Information database (www.firenorth.org.au/nafi2/, last accessed August, 2017). Burned patches were processed by the Darwin Centre for bushfires research at Charles Darwin University (for Northern Territory fire scars) and Cape York Peninsula sustainable futures (for Queensland). They were obtained from multitemporal comparison of 250 m MODIS imagery, using segmentation and visual interpretation.

Finally, as an example of temperate fires, we selected the whole state of California (409,719 km²). These fire perimeters were downloaded from the Fire and Resource Assessment program (FRAP) webpage (frap.fire.ca.gov, last accessed August 2017). These were obtained by several agencies: CAL FIRE/FRAP, the USDA Forest Service Region 5 Remote Sensing Lab, the Bureau of Land Management, and the National Park Service.

The three test sites included a total area of 2,528,471 km², with the main land covers comprising Rainfed cropland (8.22%), Tree cover, needleleaved evergreen, closed to open >15% (23.84%); Shrubland (28.33%); Grassland (8.93%); Sparse vegetation <15% (tree shrub herbaceous cover, 13.96%); and Water bodies (6.87%).

In addition to comparing the RF results with these national fire perimeters, we also cross-tabulated them with the two existing MODIS BA products: the MCD45 and the MCD64 (both v5.1), to check consistency with existing BA products. It is important to emphasize that we did not run this comparison to assess the RF models, but just to check their performance and accuracy for selected areas and periods where reliable fire information was available. Comparison with fire perimeters and global BA products

cannot be properly considered as a validation exercise, but it provides a first evaluation on the potential usability of our RF models for global BA mapping.

Comparison between both MODIS and RF products with reference fire perimeters was based on the generation of standard confusion matrices (same as Table 3). Burned perimeters from the national fire services were rasterized to 500×500 m to facilitate comparison with BA products. Omission and commission errors, overall accuracy and error balance were computed from the confusion matrices as follows:

$$\text{Omission error} = \frac{E_{+1} - E_{11}}{E_{+1}} \quad (10)$$

$$\text{Commission error} = \frac{E_{1+} - E_{11}}{E_{1+}} \quad (11)$$

$$\text{Overall accuracy} = \frac{E_{11} - E_{22}}{E_{\Sigma}} \quad (12)$$

And error balance was computed from Relative Bias (RB) following Padilla [86] as:

$$\text{Relative Bias} = \frac{E_{1+} - E_{+1}}{E_{+1}} = \frac{E_{12} - E_{21}}{E_{+1}} \quad (13)$$

All with the same meanings as in Table 3. RB indicates whether the BA product has a proper balance between omission and commission errors.

3. Results

3.1. Attribute Selection

Table 4 presents the relative importance of the different input attributes for the RF Full and RNIR models. As was expected, the distance to the closest HS and the difference in NIR reflectance were found to be the most relevant variables in both models. The importance of the SWIR reflectances was also evident in the RF full model, while GEMI performed better than other vegetation indices in the R-NIR spectral space. The auxiliary variables were found not to be relevant for the Full model, with the exception of the GFED Regions. For this reason, they were not included in the RNIR model. The final RF models had 17 attributes for the Full model and 14 for the RNIR model (Table 4).

Table 4. Ranking of each attribute in the RF Full and RNIR models. Asterisks show which variables were included in the final models. NA: Non-applicable (cannot be computed with R, NIR bands).

Band	Importance (%)		Band	Importance (%)	
	RF Full Model	RF RNIR Model		RF Full Model	RF RNIR Model
HS distance	4.54 *	14.69 *	Slope	2.00	Not included
DIF_B2	4.43 *	13.97 *	VARI_post	1.96	NA
DIF_NDWI5	3.32 *	NA	Elevation	1.94 *	8.92 *
DIF_GEMI	3.07 *	8.17 *	EVI_post	1.93	NA
B2 post	2.98 *	5.78 *	B6 post	1.93	NA
NDWI5_pre	2.90	NA	GFED Regions	1.92 *	6.99 *
DIF_MIRBI	2.87 *	NA	EVI_pre	1.90	NA
DIF_EVI	2.70	NA	SAVI_pre	1.85	4.39 *
DIF_NDWI6	2.63	NA	B6 pre	1.81	NA
MIRBI_pre	2.60 *	NA	DIF_VARI	1.78 *	NA
DIF_NBR	2.58 *	NA	NBR_post	1.72	NA
DIF_SAVI	2.44	7.21 *	B7 post	1.72	NA
MIRBI_post	2.44	NA	NBR_pre	1.71 *	NA
VARI_pre	2.42 *	NA	B7 pre	1.69	NA
GEMI_post	2.42	5.04 *	B4 post	1.68	NA
NDWI5_post	2.39 *	NA	B3 post	1.60	NA
NDWI6_pre	2.35	NA	B1post	1.53	5.89 *
B2 pre	2.33	5.27 *	B3 pre	1.46	NA
B5 post	2.25 *	NA	B4 pre	1.39	NA

Table 4. Cont.

Band	Importance (%)		Band	Importance (%)	
	RF Full Model	RF RNIR Model		RF Full Model	RF RNIR Model
B5 pre	2.24 *	NA	B1 pre	1.34	4.88 *
NDWI6_post	2.04	NA	LC_cci	1.28	Not included
SAVI_post	2.04	5.32 *	Olson	1.13	Not included
GEMI_pre	2.01 *	3.48 *	Aspect	0.74	Not included

* significant differences.

3.2. Selection of RF Parameters

Figure 3 shows the variation of balanced accuracy for several combinations of number of trees and number of attributes for the RF Full model. Even though all models had a highly balanced accuracy (>0.92), the highest values were found for those models with more parameters and greater numbers of trees, as was expected. The impact of number of trees was more evident than number of attributes, particularly when they were below 8.

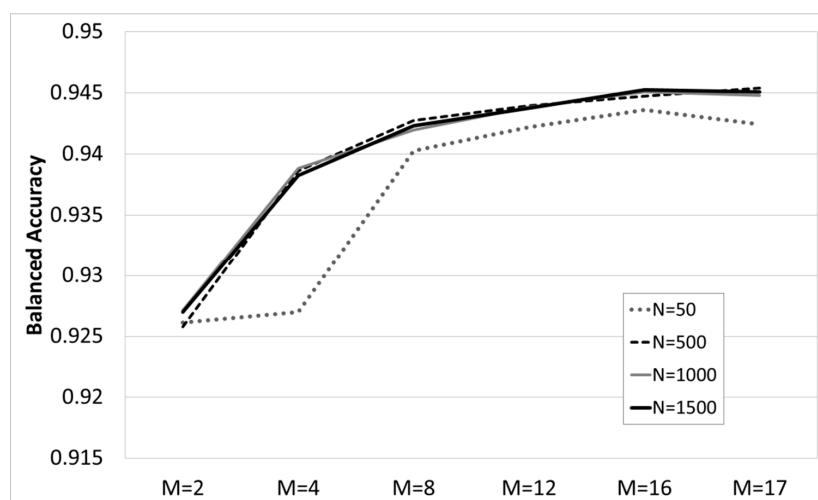


Figure 3. Variation in balanced accuracy for different N and M parameters.

Considering both accuracy and complexity, we finally chose a model with medium-high accuracy and medium-low complexity, which included $N = 600$ trees and $M = 5$ attributes. The accuracy of this RF full model for 20% of our test data (not used in the calibration of the model) provided a balanced accuracy of 0.94 with a commission error for the burned category of 0.52 and an omission error of 0.1.

The same set of parameters ($N = 600$ and $M = 5$) were used for the RF RNIR model. Using again 80% for calibration and the remaining 20% for validation of the training dataset. This model obtained a balanced accuracy of 0.93 with a commission error of 0.55 and an omission error of 0.11, similar to the RF Full model.

3.3. Classification Thresholds

The RF models generated from the training dataset were applied to classify the three study sites previously described. We tested several proportions of votes to perform the final classification of BA in the test sites. Tree assignment was tested in 10% intervals between 20% and 80%. Table 5 shows the results for the combined analysis of the three sites in 2008. For simplicity, we have included only the outputs of the RF full model, but results were similar for the RF RNIR. As was expected, omission and commission errors changed with the different thresholds, reducing omission errors when the threshold decreased, and reducing commission errors when it increased.

Table 5. Commission errors (CE), Omission errors (OE) and Relative Bias (relB) for the three test sites and different assignment thresholds. RF Full Model for 2008.

	CE	OE	relB
RF full 20%	0.35	0.11	0.38
RF full 30%	0.18	0.27	-0.11
RF full 40%	0.22	0.26	-0.05
RF full 50%	0.17	0.36	-0.23
RF full 60%	0.12	0.50	-0.43
RF full 70%	0.08	0.67	-0.63
RF full 80%	0.06	0.84	-0.83

The best balance between omission and commission errors was found at the 40% threshold. This is, to a certain extent, reasonable, as on one hand it is a balanced value, and on the other tends to take into better account the lower prevalence of burned over unburned pixels than the standard 50% value. For this reason, we have used this threshold for the other years, as well as for the RF RNIR model.

3.4. RF Classification Results

Figures 4–6 include results of the RF full model for the three areas. The results of the RF RNIR model were quite similar in terms of geographical distribution of burned patches, and therefore we have not included them to avoid redundancy.

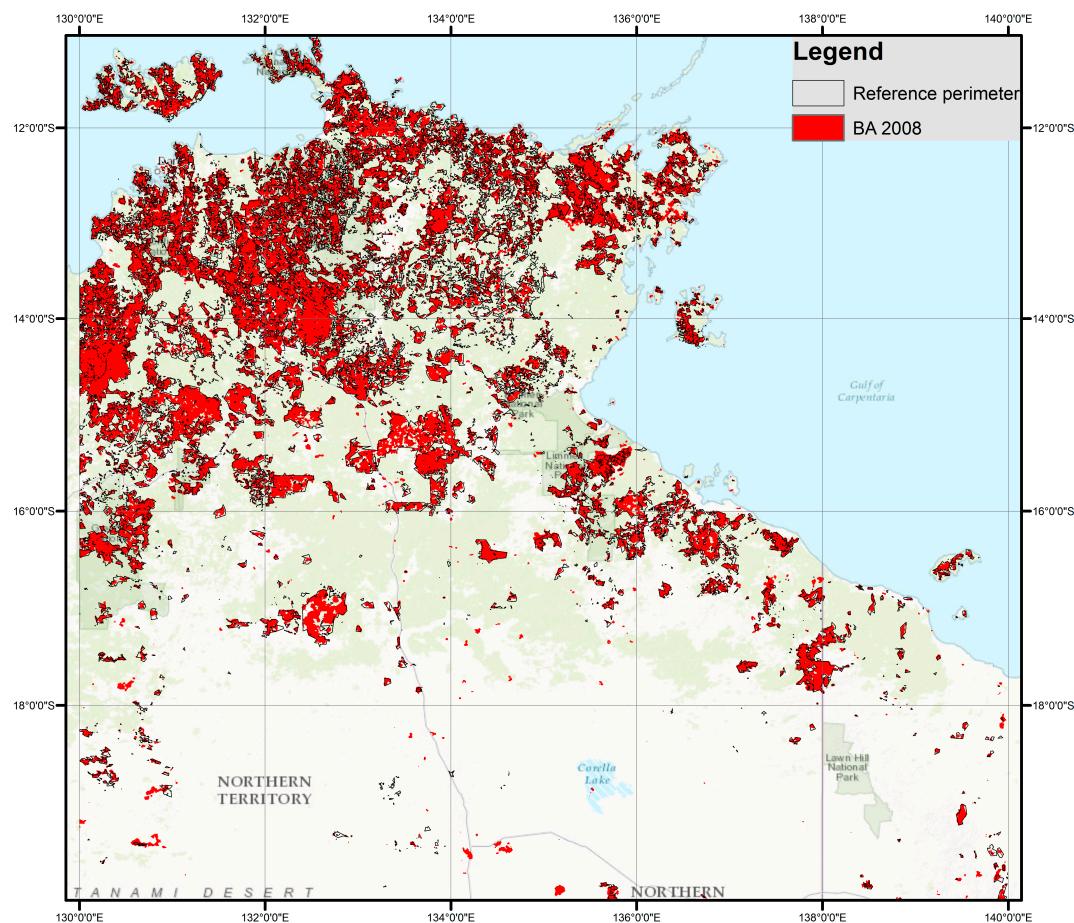


Figure 4. Burned areas detected by the RF full model in Australia (2008).

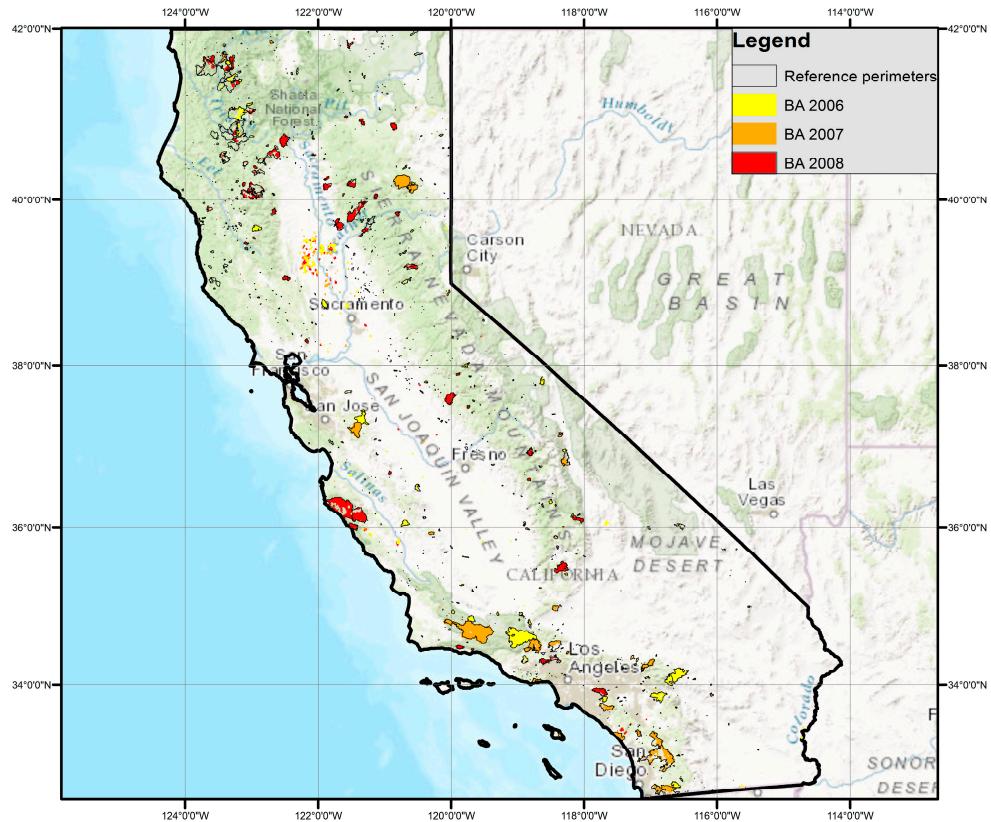


Figure 5. RF full model results for 2006 to 2008 in California.

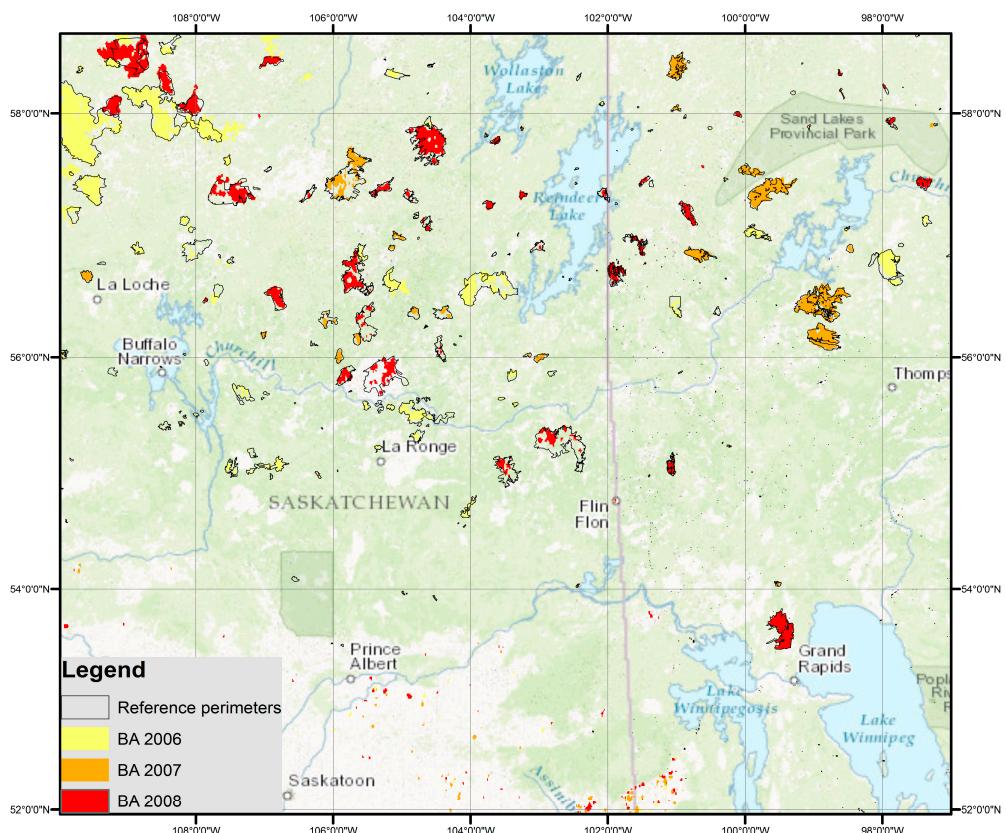


Figure 6. RF full model results for 2006 to 2008 in Canada.

In accordance with data obtained from fire perimeters, more than 661,100 km² were burned in the three sites over the three years. Annual BA was much lower in 2008, with 140,200 km², while 2006 and 2007 had similar fire incidences (242,700 and 278,100 km², respectively). The vast majority of that BA occurred in Australia, totaling 93.89% of all burned area in the three sites and years. The lower BA value of 2008 was caused by the lower fire occurrence of Australia, where BA decreased to 122,600 km² from 228,600 and 269,700 km² for 2006 and 2007, respectively. California had a total 12,700 km² burned in the three years, with greater occurrence in 2008 (more than 5540 km²) and lower in 2006 (close to 3000 km²). Canada totaled more than 27,700 km² burned, with much higher occurrence in 2008 and 2006 (both more than 10,000 km²) than 2007 (4520 km² burned).

Results from the two RF models present similar total BA values to those obtained from fire perimeters. The RF Full model estimates the total BA in 603,000 km² and the RNIR model in 651,900 km². The former had greater underestimation than the latter, particularly at the Australian site (566,900 km² and 600,000 km², respectively, versus the 620,000 km² calculated from fire perimeters). The RF Full model provided a closer estimation to the reference data than the RNIR model for California (10,900 km² and 18,300 versus the reference of 12,700), as well as for the Canadian site (25,200 km² Full model, 33,500 km² RNIR model and 27,700 km² reference data).

Figure 4 shows the results of the RF Full model for the Australian site, along with the reference perimeters. The figure shows only the 2008 results, since several areas were burned more than once in the three years. The region closer to the coastline had more fires (and more persistent ones), while the Southern part burned less and less frequently, as it had a smaller amount of available fuel. Most of the observed omissions of the RF Full model were pixels within burned patches. Commissions were found in areas close to actual burn perimeters, probably in transitional areas with partial burns.

Figure 5 shows the spatial distribution of classified BA in California. The greatest disagreement with our reference perimeters was found in the Central Valley (Sacramento area). These were associated with agriculture areas with low proportions of natural vegetation. Since our reference perimeters were derived by different forest agencies, agricultural burnings were usually absent, and therefore these apparent commission errors may in fact be real burnings. Most of the omission errors were found in the northern fringe, in forested areas with higher moisture than the southern area, and lower fire sizes. The area from Santa Barbara to San Diego concentrates the largest burned areas with a mixture of forest and shrub land. They were in agreement with the reference perimeters.

Finally, Figure 6 shows the results of the Canadian site. Similar to California, the most obvious disagreements with reference perimeters were found in the agricultural areas, located in the southern region of the image. Again, it is important to remember that agricultural fires are not included in the LFDB, which is restricted to wildland areas. The omission errors were associated to patches within large fires in the northern area, where RF outputs did not properly identify the complete burned patch.

A quantitative comparison of RF model outputs and reference perimeters is displayed in Tables 6 and 7, showing the omission and commission errors (or, more precisely, the disagreements with national fire databases) for the three study sites and the three classification years, for both the RF Full and RNIR models. Overall accuracies are not included in the tables, but values were always above 0.91. For the Australian site, the RF model showed OA > 0.91, with higher values for 2008 and 2006, when more area was burned. Both Full and RNIR models performed similarly. Canada and California had much higher OA (all above 0.99 in both RF models), due to their having a much lower proportion of BA than Australia, and the higher accuracy of the models at predicting unburned pixels. The mean and total values (summing up the three sites) consider the actual BA occurring in each study site and year, with a clear predominance of the results obtained in the Australian area, as it had the vast majority of total BA.

The RF Full model results in Table 6 show that the average commissions and omissions for all three years and at all three sites were 0.16 and 0.23, respectively. The commission errors were lower at all sites and years than the omission errors, but they can be considered to be well balanced. Omission errors were higher for 2008, particularly in California.

Table 6. RF Full model classification results.

Year	Commission Errors				Omission Errors			
	Australia	California	Canada	Global	Australia	California	Canada	Global
2006	0.13	0.34	0.30	0.14	0.21	0.32	0.33	0.22
2007	0.15	0.24	0.32	0.15	0.24	0.25	0.32	0.24
2008	0.22	0.23	0.23	0.22	0.25	0.50	0.38	0.26
Mean	0.16	0.27	0.28	0.16	0.23	0.38	0.35	0.23

Table 7. RF RNIR model classification results.

	Commission Errors				Omission Errors			
	Australia	California	Canada	Total	Australia	California	Canada	Total
2006	0.16	0.56	0.38	0.16	0.19	0.18	0.23	0.19
2007	0.18	0.45	0.45	0.19	0.22	0.16	0.22	0.22
2008	0.25	0.41	0.32	0.26	0.23	0.34	0.27	0.24
Mean	0.19	0.47	0.37	0.19	0.21	0.25	0.25	0.21

Table 7 shows the omission and commission errors for RF RNIR model. Commission errors were generally higher and omissions lower than for RF Full model. In California, commission errors were 20% higher over the three years than the Full model, while in Canada and Australia the values were only slightly higher (<10%). In terms of omission errors, the RNIR model performed slightly better than the Full model in Australia and Canada, and showed about 15% better results than the Full model in California.

3.5. Comparison with Existing MODIS BA Products

Results from the RF models were compared with existing MODIS BA products and reference fire perimeters for the three test sites (Table 8). To summarize the presentation of the results, errors and relative bias were computed by combining the three years. Commission errors of MCD45 and MCD64 products tended to be lower than for the RF models, while omission errors were commonly higher, particularly for the MCD45 product. The RF Full model was more balanced than MODIS products for Australia and California, and only slightly lower than MCD64 for Canada. The RF RNIR model was very well balanced in Australia, but it provided higher commission errors in Canada and California than the other products, and it tended to overestimate in these two sites.

Table 8. Mean omission and commission errors and relative bias (2006 to 2008).

	Australia			Canada			California			Total		
	CE	OE	relB	CE	OE	relB	CE	OE	relB	CE	OE	relB
MCD45	0.09	0.32	-0.26	0.30	0.57	-0.38	0.35	0.55	-0.31	0.09	0.33	-0.26
MCD64	0.09	0.29	-0.21	0.21	0.28	-0.09	0.12	0.29	-0.19	0.10	0.29	-0.21
RF full	0.16	0.23	-0.09	0.28	0.35	-0.10	0.27	0.38	-0.15	0.16	0.23	-0.09
RF RNIR	0.19	0.21	-0.03	0.37	0.25	0.19	0.47	0.25	0.43	0.19	0.21	-0.03

Both RF models showed higher commission errors than the MODIS products, although still lower than 0.3 in the case of RF Full model (The RNIR model provided higher CE). In terms of omission errors, the MCD45 showed high errors for California and Canada (>0.55), while the MCD64 had values close to 0.3. The RF full model showed lower omission errors than MCD45, and slightly higher omission errors than MCD64 for California and Canada. The RNIR model had the lowest omission errors at both sites. Similar trends were observed for all three years.

4. Discussion

This paper has presented the generation and first assessment of a BA classifier based on two RF models. The models were trained using a statistically designed sample of reference burned areas generated from Landsat multitemporal images, covering the global variety of fire characteristics. Therefore, as the RF models were trained and internally assessed (using OOB statistics) with a global sample, they can be considered suitable for global mapping of burned areas.

We should emphasize that our results have not been validated in a statistically rigorous way, which would require global process and reference BA datasets [72,87]. The alternative was to analyze the accuracy of the algorithm at selected representative sites and compare the results with existing reference BA data. We selected three test sites located in the main ecosystems affected by fire occurrence: Tropical, Temperate and Boreal, where fire perimeters were available. Comparison of RF results with these reference BA datasets showed high overall accuracy values (>0.9), although the burned category had omission and commission errors of around 0.25. The similar accuracy of both models evidences interesting potentials to apply analogous BA algorithms to sensors that do not acquire SWIR bands, such as NOAA-AVHRR, ENVISAT-MERIS, or Sentinel-3 OLCI, although they would still require the location of active fires. Comparison with standard MODIS BA products provided higher commission, but lower errors, for the RF models. Therefore, both RF results showed lower error bias than MODIS products, particularly in comparison with the MCD45. Lower errors were found in Australia, which had much larger BAs than California and Canada.

Disagreements found between reference BA and RF results may also be caused by some problems of the former datasets. For instance, fires from California and Canada may not include agricultural burnings, while they are detected as burned in both the RF and the standard NASA BA products. Therefore, a share of the commission errors reported may be in fact real burnings occurring on crop residues. In terms of omission errors, the main problem of the reference datasets is the allocation of unburned islands within burned patches, which is not done systematically [88]. Another potential problem may arise from the consideration of small fires ($<1 \text{ km}^2$), which are included in reference perimeters, but are undetectable from $500 \times 500 \text{ m}$ pixels. In spite of these limitations, our results show at least a first assessment of the RF models, as they include very extended areas ($>2.5 \text{ million km}^2$) in different years and diverse ecoregions.

In terms of the limitations of our RF models, several aspects should be considered:

1. Temporal generalization. Several authors have pointed out the potential problems of training RF classifiers with a single year and then applying them to other years [89]. In our case, the models were trained for 2008 and tested in two other years. Results showed similar accuracy for the three years, which provides good support for using the RF models to generate BA time series analyses.
2. Spatial generalization. Our RF models were not trained from data extracted from the three test sites, as it is a common practice of many RF classifiers [24,36,90], but rather from an independent dataset that was selected using a statistical sample design to include the diversity of burning conditions worldwide (at least for 2008, the year of the reference perimeters), as it was originally designed by validating global BA products [72]. For this reason, the resultant RF models should be applicable globally without further calibration, assuming 2008 burned conditions are representative of global fire patterns. In this regard, it could be considered a global automatic classification. We have shown this with three very different sites and years, fully independent from our training dataset.
3. Sensor generalization. RF has been recognized in very diverse applications as a flexible approach to integrating different sensors and scales [25,26,35,79]. This capacity makes it possible to extend a similar concept of a RF-based BA algorithm to other sensors such as VIIRS (on board the NPOES), OLCI or SLSTR (on board Sentinel-3), or even for classifying historical datasets, such as those acquired by the AVHRR (on board NOAA from 1979). However, the dependency of the

- active fire information that the RF models (as well as the MCD64 product) show needs to be addressed, since active fire detections prior to the MODIS era are quite unreliable.
- 4. Time constraints. Regarding the performance of RF models, the computational cost of training a RF model is low in comparison with other machine-learning algorithms like support vector machines (SVM) [91]. The processing time depends on the complexity of the model, and the number of trees and attributes. Our models required about three hours for training and 4 min for classifying each 1200×1200 km image (using a 24 core computer with 48 Gb of RAM).
 - 5. Training phase. The strengths of our training rely on the global significance of the sample, but it implied a large effort to generate the training dataset, as data from more than 15,000 images had to be retrieved. Since globally speaking fire occurrence is an unusual event, the sample was very unbalanced, with only 0.76% of BA in the total area covered by Landsat images. Several strategies may be used to cope with this problem, including under and oversampling of the target categories [92,93]. In our case, every tree was forced to include at least 10% of burned pixels [32,77,78]. This ensured the inclusion of all pixels detected as burned in the ensemble of trees. Another possibility of RF is to estimate the proportion of burned area, but this estimation also requires a balanced database, because regression trees are also affected by this issue. Mapping proportions is certainly interesting and feasible at regional scales through fuzzy classifications or spectral mixture models [94], but is very difficult to implement at a global scale. Future improvements of the RF models may include additional training datasets, acquired from different years, thus extending the characterization of burn patterns to the impact of climatic cycles.
 - 6. Attribute selection. As may be expected, the most important attributes were those related to the distance from hotspots, post-fire NIR reflectance and temporal differences in NIR reflectance. In the Full model, contributions of the SWIR bands were also relevant, following the well-known spectral sensitivity of this region for characterizing post-fire conditions [11,19,56,95]. The predominance of distance to HS is also well known in BA mapping, as many authors have emphasized the importance of active fire information for better discriminating burned pixels. This evidence has led to the generation of hybrid algorithms [61,62], which are currently the state of the art in global BA mapping [19,21]. Following HS distance, we found that post-fire NIR and changes in NIR reflectance were the most explicative variables. Again, several authors have shown that NIR reflectance exhibits the highest sensitivity for detecting changes caused by burnings. This could be related to the removal of the chlorophyll content and water, as the leaf is scorched, as well as the decrease in leaf area index [73].
 - 7. RF parameters. Different combinations of both the number of trees and the number of attributes were tested. The final RF models balanced performance and accuracy. A large number of trees and attributes implied a severe increase in the classification time, which reduced the performance of the model. Since we classified daily images, the yearly BA requires running the RF models for each site 365 times (with N classifications for each pixel, with N being the number of trees). Complex models implied a severe computational burden, and would make the final model for classifying global BA coverage unrealistic. The compromise between complexity and accuracy of the model was obtained through the analysis of accuracy changes with model complexity (Figure 3). The actual attributes were selected by analyzing their relative importance in the discrimination of burned area and their mutual correlation. This is a common strategy, as redundant variables may bias the generation of the RF models [96,97]. The selection of the attributes provided by RF may also be obtained through other techniques widely used in data mining [98,99].

5. Conclusions

This study presents a methodology for generating a burned area algorithm using the RF classifier. Input data were the MODIS BRDF corrected reflectances, active fire information, and auxiliary geographical coverages. The model was trained from a set of 130 statistically selected sampled sites,

and was later applied to three large test sites located in Australia, Canada and California. The most relevant variables for BA classification were post-fire NIR, NIR change and distance to active fires. The SWIR bands were not found to be as important as other studies had shown, since the performance of the RF Full model did not significantly change from the RF RNIR model.

RF was proved to be a robust classifier of BA, showing similar accuracy in different sites and years, with lower BA errors in Tropical than Boreal and Temperate sites. Omission and commission errors were estimated by comparing RF results with national fire perimeters. RF outputs generally had lower omission and higher commission errors than existing MODIS BA products. They also provided a more balanced accuracy. The RF models showed great potential to be applied to other sensors and periods in order to generate consistent time series of BA data, which are in great demand among the climate modeling community.

Acknowledgments: This study has been undertaken within the ESA Fire_cci project, from which images and computer resources were acquired. Ruben Ramo is being funded by the University of Alcalá's Pre-doctoral Fellowship Program.

Author Contributions: Rubén Ramo and Emilio Chuvieco designed the study, Emilio Chuvieco participates in the discussion of the different phases and prepares the writing of the manuscript, Ruben Ramo performed the RF analysis and processed the training and classification analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kloster, S.; Mahowald, N.; Randerson, J.; Lawrence, P. The impacts of climate, land use, and demography on fires during the 21st century simulated by clm-cn. *Biogeosciences* **2012**, *9*, 509–525. [[CrossRef](#)]
2. Thonicke, K.; Spessa, A.; Prentice, I.C.; Harrison, S.P.; Dong, L.; Carmona-Moreno, C. The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: Results from a process-based model. *Biogeosciences* **2010**, *7*, 1991–2011. [[CrossRef](#)]
3. Van der Werf, G.R.; Randerson, J.T.; Giglio, L.; Collatz, G.; Mu, M.; Kasibhatla, P.S.; Morton, D.C.; DeFries, R.S.; Jin, Y.; van Leeuwen, T.T. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys.* **2010**, *10*, 11707–11735. [[CrossRef](#)]
4. Schoennagel, T.; Nelson, C.R.; Theobald, D.M.; Carnwath, G.C.; Chapman, T.B. Implementation of national fire plan treatments near the wildland–urban interface in the western united states. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10706–10711. [[CrossRef](#)] [[PubMed](#)]
5. Global Climate Observing System (GCOS). *Guideline for the Generation of Datasets and Products Meeting GCOS Requirements*; World Meteorological Organization: Geneva, Switzerland, 2010.
6. Hollmann, R.; Merchant, C.J.; Saunders, R.W.; Downy, C.; Buchwitz, M.; Cazenave, A.; Chuvieco, E.; Defourny, P.; Leeuw, G.D.; Forsberg, R.; et al. The esa climate change initiative: Satellite data records for essential climate variables. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1541–1552. [[CrossRef](#)]
7. Mouillot, F.; Schultz, M.G.; Yue, C.; Cadule, P.; Tansey, K.; Ciais, P.; Chuvieco, E. 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.* **2014**, *26*, 64–79. [[CrossRef](#)]
8. Chuvieco, E.; Yue, C.; Heil, A.; Mouillot, F.; Alonso-Canas, I.; Padilla, M.; Pereira, J.M.; Oom, D.; Tansey, K. A new global burned area product for climate assessment of fire impacts. *Glob. Ecol. Biogeogr.* **2016**, *25*, 619–629. [[CrossRef](#)]
9. Riaño, D.; Moreno Ruiz, J.; Isidoro, D.; Ustin, S. Global spatial patterns and temporal trends of burned area between 1981 and 2000 using noaa-nasa pathfinder. *Glob. Chang. Biol.* **2007**, *13*, 40–50. [[CrossRef](#)]
10. Tansey, K.; Grégoire, J.M.; Defourny, P.; Leigh, R.; Peckel, J.F.; Bogaert, E.V.; Bartholome, J.E. A new, global, multi-annual (2000–2007) burnt area product at 1 km resolution. *Geophys. Res. Lett.* **2008**, *35*, L01401. [[CrossRef](#)]
11. Roy, D.; Jin, Y.; Lewis, P.; Justice, C. Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sens. Environ.* **2005**, *97*, 137–162. [[CrossRef](#)]

12. Bastarrika, A.; Chuvieco, E.; Martín, M.P. Mapping burned areas from landsat tm/etm+ data with a two-phase algorithm: Balancing omission and commission errors. *Remote Sens. Environ.* **2011**, *115*, 1003–1012. [[CrossRef](#)]
13. Pu, R.; Gong, P. Determination of burnt scars using logistic regression and neural network techniques from a single post-fire landsat-7 etm+ image. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 841–850. [[CrossRef](#)]
14. Koutsias, N.; Karteris, M. Burned area mapping using logistic regression modeling of a single post-fire landsat-5 thematic mapper image. *Int. J. Remote Sens.* **2000**, *21*, 673–687. [[CrossRef](#)]
15. Kachmar, M.; Sanchez-Azofeifa, G.A. Detection of post-fire residuals using high- and medium-resolution satellite imagery. *For. Chron.* **2006**, *82*, 177–186. [[CrossRef](#)]
16. Mitri, G.H.; Gitas, I.Z. Mapping post-fire forest regeneration and vegetation recovery using a combination of very high spatial resolution and hyperspectral satellite imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *20*, 60–66. [[CrossRef](#)]
17. Van Leeuwen, T.T.; Peters, W.; Krol, M.C.; van der Werf, G.R. Dynamic biomass burning emission factors and their impact on atmospheric co mixing ratios. *J. Geophys. Res. Atmos.* **2013**, *118*, 6797–6815. [[CrossRef](#)]
18. Giglio, L.; Randerson, J.T.; Werf, G.R. Analysis of daily, monthly, and annual burned area using the fourth generation global fire emissions database (gfed4). *J. Geophys. Res. Biogeosci.* **2013**, *118*, 317–328. [[CrossRef](#)]
19. Giglio, L.; Loboda, T.; Roy, D.P.; Quayle, B.; Justice, C.O. An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sens. Environ.* **2009**, *113*, 408–420. [[CrossRef](#)]
20. Plummer, S.; Arino, O.; Ranera, F.; Tansey, K.; Chen, J.; Dedieu, G.; Eva, H.; Piccolini, I.; Leigh, R.; Borstlap, G. The globcarbon initiative: Multi-sensor estimation of global biophysical products for global terrestrial carbon studies. In Proceedings of the Envisat & ERS Symposium, Salzburg, Austria, 6–10 September 2004.
21. Alonso-Canas, I.; Chuvieco, E. Global burned area mapping from envisat-meris data. *Remote Sens. Environ.* **2015**, *163*, 140–152. [[CrossRef](#)]
22. Chang, D.; Song, Y. Comparison of L3JRC and MODIS global burned area products from 2000 to 2007. *J. Geophys. Res.* **2009**, *114*. [[CrossRef](#)]
23. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
24. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sánchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [[CrossRef](#)]
25. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
26. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [[CrossRef](#)]
27. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [[CrossRef](#)]
28. Pelletier, C.; Valero, S.; Ingla, J.; Champion, N.; Dedieu, G. Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. *Remote Sens. Environ.* **2016**, *187*, 156–168. [[CrossRef](#)]
29. Immitzer, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-band worldview-2 satellite data. *Remote Sens.* **2012**, *4*, 2661–2693. [[CrossRef](#)]
30. Lawrence, R.L.; Wood, S.D.; Sheley, R.L. Mapping invasive plants using hyperspectral imagery and breiman cutler classifications (randomforest). *Remote Sens. Environ.* **2006**, *100*, 356–362. [[CrossRef](#)]
31. Ok, A.O.; Akar, O.; Gungor, O. Evaluation of random forest method for agricultural crop classification. *Eur. J. Remote Sens.* **2012**, *45*, 421–432. [[CrossRef](#)]
32. Han, H.; Wang, W.-Y.; Mao, B.-H. Borderline-smote: A new over-sampling method in imbalanced data sets learning. In Proceedings of the International Conference on Intelligent Computing, Hefei, China, 23–26 August 2005; Springer: Berlin/Heidelberg, Germany, 2005; pp. 878–887.
33. Pommerening, A. Approaches to quantifying forest structures. *Forestry* **2002**, *75*, 305–324. [[CrossRef](#)]
34. Mutanga, O.; Adam, E.; Cho, M.A. High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 399–406. [[CrossRef](#)]

35. Naidoo, L.; Cho, M.; Mathieu, R.; Asner, G. 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.* **2012**, *69*, 167–179. [[CrossRef](#)]
36. Oliveira, S.; Oehler, F.; San-Miguel-Ayanz, J.; Camia, A.; Pereira, J.M. Modeling spatial patterns of fire occurrence in mediterranean europe using multiple regression and random forest. *For. Ecol. Manag.* **2012**, *275*, 117–129. [[CrossRef](#)]
37. Holden, Z.A.; Morgan, P.; Evans, J.S. A predictive model of burn severity based on 20-year satellite-inferred burn severity data in a large southwestern us wilderness area. *For. Ecol. Manag.* **2009**, *258*, 2399–2406. [[CrossRef](#)]
38. Archibald, S.A.; Roy, D.P.; Van Wilgen, B.W.; Scholes, R.J. What limits fire? An examination of drivers of burnt area in southern africa. *Glob. Chang. Biol.* **2009**, *15*, 613–630. [[CrossRef](#)]
39. Aldersley, A.; Murray, S.J.; Cornell, S.E. Global and regional analysis of climate and human drivers of wildfire. *Sci. Total Environ.* **2011**, *409*, 3472–3481. [[CrossRef](#)] [[PubMed](#)]
40. Chuvieco, E.; Riano, D.; Aguado, I.; Cocco, D. Estimation of fuel moisture content from multitemporal analysis of landsat thematic mapper reflectance data: Applications in fire danger assessment. *Int. J. Remote Sens.* **2002**, *23*, 2145–2162. [[CrossRef](#)]
41. Trigg, S.; Flasse, S. An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah. *Int. J. Remote Sens.* **2001**, *22*, 2641–2647. [[CrossRef](#)]
42. Schaaf, C.B.; Gao, F.; Strahler, A.H.; Lucht, W.; Li, X.; Tsang, T.; Strugnell, N.C.; Zhang, X.; Jin, Y.; Muller, J.-P.; et al. First operational brdf, albedo nadir reflectance products from MODIS. *Remote Sens. Environ.* **2002**, *83*, 135–148. [[CrossRef](#)]
43. Liu, Z.; Wimberly, M.C.; Dwomoh, F.K. Vegetation dynamics in the upper guinean forest region of west africa from 2001 to 2015. *Remote Sens.* **2016**, *9*, 5. [[CrossRef](#)]
44. Huete, A.R. A soil-adjusted vegetation index (savi). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
45. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
46. Garcia, M.; Chuvieco, E. Assessment of the potential of sac-c/mmrs imagery for mapping burned areas in spain. *Remote Sens. Environ.* **2004**, *92*, 414–423. [[CrossRef](#)]
47. Chuvieco, E.; Martín, M.P.; Palacios, A. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *Int. J. Remote Sens.* **2002**, *23*, 5103–5110. [[CrossRef](#)]
48. Stroppiana, D.; Bordogna, G.; Carrara, P.; Boschetti, M.; Boschetti, L.; Brivio, P. A method for extracting burned areas from landsat tm/etm+ images by soft aggregation of multiple spectral indices and a region growing algorithm. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 88–102. [[CrossRef](#)]
49. Pinty, B.; Verstraete, M.M. Gemi: A non-linear index to monitor global vegetation from satellites. *Vegetatio* **1992**, *101*, 15–20. [[CrossRef](#)]
50. Barbosa, P.M.; Grégoire, J.M.; Pereira, J.M.C. An algorithm for extracting burned areas from time series of avhrr gac data applied at a continental scale. *Remote Sens. Environ.* **1999**, *69*, 253–263. [[CrossRef](#)]
51. López García, M.J.; Caselles, V. Mapping burns and natural reforestation using thematic mapper data. *Geocarto Int.* **1991**, *1*, 31–37. [[CrossRef](#)]
52. Brewer, C.K.; Winne, J.C.; Redmond, R.L.; Opitz, D.W.; Mangrich, M.V. Classifying and mapping wildfire severity: A comparison of methods. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 1311–1320. [[CrossRef](#)]
53. Rogan, J.; Franklin, J. Mapping wildfire burn severity in southern california forests and shrublands using enhanced thematic mapper imagery. *Geocarto Int.* **2001**, *16*, 89–99. [[CrossRef](#)]
54. Gao, B.C. Ndwi: A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [[CrossRef](#)]
55. Chuvieco, E.; Riaño, D.; Danson, F.M.; Martín, M.P. Use of a radiative transfer model to simulate the post-fire spectral response to burn severity. *J. Geophys. Res. Biosci.* **2006**, *111*. [[CrossRef](#)]
56. Gitelson, A.A.; Stark, R.; Grits, U.; Rundquist, D.; Kaufman, Y.; Derry, D. Vegetation and soil lines in visible spectral space: A concept and technique for remote estimation of vegetation fraction. *Int. J. Remote Sens.* **2002**, *23*, 2537–2562. [[CrossRef](#)]
57. Gitelson, A.; Kaufman, J.Y.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [[CrossRef](#)]

58. Schneider, P.; Roberts, D.; Kyriakidis, P. A vari-based relative greenness from MODIS data for computing the fire potential index. *Remote Sens. Environ.* **2008**, *112*, 1151–1167. [[CrossRef](#)]
59. Jin, Y.; Randerson, J.T.; Goetz, S.J.; Beck, P.S.; Loranty, M.M.; Goulden, M.L. The influence of burn severity on postfire vegetation recovery and albedo change during early succession in north american boreal forests. *J. Geophys. Res. Biogeosci.* **2012**, *117*, G01036. [[CrossRef](#)]
60. Bastarrika, A.; Alvarado, M.; Artano, K.; Martinez, M.; Mesanza, A.; Torre, L.; Ramo, R.; Chuvieco, E. Bams: A tool for supervised burned area mapping using landsat data. *Remote Sens.* **2014**, *6*, 12360–12380. [[CrossRef](#)]
61. Roy, D.P.; Giglio, L.; Kendall, J.D.; Justice, C.O. Multi-temporal active-fire based burn scar detection algorithm. *Int. J. Remote Sens.* **1999**, *20*, 1031–1038. [[CrossRef](#)]
62. Fraser, R.H.; Li, Z.; Cihlar, J. Hotspot and ndvi differencing synergy (hands): A new technique for burned area mapping over boreal forest. *Remote Sens. Environ.* **2000**, *74*, 362–376. [[CrossRef](#)]
63. Hantson, S.; Arneth, A.; Harrison, S.P.; Kelley, D.I.; Prentice, I.C.; Rabin, S.S.; Archibald, S.; Mouillet, F.; Arnold, S.R.; Artaxo, P. The status and challenge of global fire modelling. *Biogeosciences* **2016**, *13*, 3359–3375. [[CrossRef](#)]
64. Hantson, S.; Pueyo, S.; Chuvieco, E. Global fire size distribution is driven by human impact and climate. *Glob. Ecol. Biogeogr.* **2015**, *24*, 77–86. [[CrossRef](#)]
65. Danielson, J.J.; Gesch, D.B. *Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010)*; US Geological Survey: Riston, FL, USA, 2011.
66. Burrough, P.A.; McDonnell, R.; McDonnell, R.A.; Lloyd, C.D. *Principles of Geographical Information Systems*; Oxford University Press: Oxford, UK, 2015.
67. Defourny, P.; Kirches, G.; Brockmann, C.; Boettcher, M.; Peters, M.; Bontemps, S.; Lamarche, C.; Schlerf, M.; Santoro, M. Land COVER CCI. Product User Guide Version 2.0. 2012. Available online: https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf (accessed on 7 October 2017).
68. Herold, M.; Woodcock, C.E.; Di Gregorio, A.; Mayaux, P.; Belward, A.S.; Latham, J.; Schmullius, C.C. A joint initiative for harmonization and validation of land cover datasets. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1719–1727. [[CrossRef](#)]
69. Olson, D.M.; Dinerstein, E.; Wikramanayake, E.D.; Burgess, N.D.; Powell, G.V.N.; Underwood, E.C.; D’amicco, J.A.; Itoua, I.; Strand, H.E.; Morrison, J.C.; et al. Terrestrial ecoregions of the world: A new map of life on earth. *BioScience* **2001**, *51*, 933–938. [[CrossRef](#)]
70. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; CRC Press: Florida, FL, USA, 1984.
71. Bylander, T. Estimating generalization error on two-class datasets using out-of-bag estimates. *Mach. Learn.* **2002**, *48*, 287–297. [[CrossRef](#)]
72. Padilla, M.; Stehman, S.V.; Chuvieco, E. Validation of the 2008 MODIS-mcd45 global burned area product using stratified random sampling. *Remote Sens. Environ.* **2014**, *144*, 187–196. [[CrossRef](#)]
73. Pereira, J.M.C. A comparative evaluation of noaa/avhrr vegetation indexes for burned surface detection and mapping. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 217–226. [[CrossRef](#)]
74. Clark, M.L.; Aide, T.M.; Grau, H.R.; Riner, G. A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the dry chaco ecoregion of south america. *Remote Sens. Environ.* **2010**, *114*, 2816–2832. [[CrossRef](#)]
75. Ayala-Izurieta, J.E.; Márquez, C.O.; García, V.J.; Recalde-Moreno, C.G.; Rodríguez-Llerena, M.V.; Damián-Carrión, D.A. Land cover classification in an ecuadorian mountain geosystem using a random forest classifier, spectral vegetation indices, and ancillary geographic data. *Geosciences* **2017**, *7*, 34. [[CrossRef](#)]
76. Khoshgoftaar, T.M.; Golawala, M.; Van Hulse, J. An empirical study of learning from imbalanced data using random forest. In Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), Patras, Greece, 29–31 October 2007; pp. 310–317.
77. Breiman, L.; Chen, C.; Liaw, A. Using random forest to learn imbalanced data. *J. Mach. Learn. Res.* **2004**, *1*–12.
78. Colditz, R.R. An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms. *Remote Sens.* **2015**, *7*, 9655–9681. [[CrossRef](#)]
79. Mellor, A.; Boukir, S.; Haywood, A.; Jones, S. Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 155–168. [[CrossRef](#)]

80. Zhou, L.; Wang, H. Loan default prediction on large imbalanced data using random forests. *Indones. J. Electr. Eng.* **2012**, *10*, 1519–1525. [[CrossRef](#)]
81. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 74–77.
82. Burton, P.J.; Parisien, M.-A.; Hicke, J.A.; Hall, R.J.; Freeburn, J.T. Large fires as agents of ecological diversity in the north american boreal forest. *Int. J. Wildland Fire* **2009**, *17*, 754–767. [[CrossRef](#)]
83. Parisien, M.A.; Peters, V.S.; Wang, Y.H.; Little, J.M.; Bosch, E.M.; Stocks, B.J. Spatial patterns of forest fires in canada, 1980–1999. *Int. J. Wildland Fire* **2006**, *15*, 361–374. [[CrossRef](#)]
84. Stocks, B.; Mason, J.; Todd, J.; Bosch, E.; Wotton, B.; Amiro, B.; Flannigan, M.; Hirsch, K.; Logan, K.; Martell, D. Large forest fires in canada, 1959–1997. *J. Geophys. Res. Atmos.* **2002**. [[CrossRef](#)]
85. Amiro, B.D.; Todd, J.B.; Wotton, B.M.; Logan, K.A.; Flannigan, M.D.; Stocks, B.J.; Mason, J.A.; Martell, D.L.; Hirsch, K.G. Direct carbon emissions from canadian forest fires, 1959–1999. *Can. J. For. Res.* **2001**, *31*, 512–525. [[CrossRef](#)]
86. Padilla, M.; Stehman, S.V.; Hantson, S.; Oliva, P.; Alonso-Canas, I.; Bradley, A.; Tansey, K.; Mota, B.; Pereira, J.M.; Chuvieco, E. Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation. *Remote Sens. Environ.* **2015**, *160*, 114–121. [[CrossRef](#)]
87. Padilla, M.; Olofsson, P.; Stehman, S.V.; Tansey, K.; Chuvieco, E. Stratification and sample allocation for reference burned area data. *Remote Sens. Environ.* **2017**. [[CrossRef](#)]
88. Boschetti, L.; Roy, D.P.; Justice, C.O.; Humber, M.L. MODIS–landsat fusion for large area 30m burned area mapping. *Remote Sens. Environ.* **2015**, *161*, 27–42. [[CrossRef](#)]
89. Malmström, C.M.; Thompson, M.V.; Juday, G.P.; Los, S.O.; Randerson, J.T.; Field, C.B. Interannual variation in global-scale net primary production: Testing model estimates. *Glob. Biogeochem. Cycles* **1997**, *11*, 367–392. [[CrossRef](#)]
90. Chan, J.C.-W.; Paelinckx, D. Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sens. Environ.* **2008**, *112*, 2999–3011. [[CrossRef](#)]
91. Mallinis, G.; Koutsias, N. Comparing ten classification methods for burned area mapping in a mediterranean environment using landsat tm satellite data. *Int. J. Remote Sens.* **2012**, *33*, 4408–4433. [[CrossRef](#)]
92. Chawla, N.V. Data mining for imbalanced datasets: An overview. In *Data Mining and Knowledge Discovery Handbook*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 853–867.
93. Liu, H.; Motoda, H. *Instance Selection and Construction for Data Mining*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013; Volume 608.
94. Quintano, C.; Fernández-Manso, A.; Roberts, D.A. Multiple endmember spectral mixture analysis (mesma) to map burn severity levels from landsat images in mediterranean countries. *Remote Sens. Environ.* **2013**, *136*, 76–88. [[CrossRef](#)]
95. Pereira, J.M.C.; Mota, B.; Privette, J.L.; Caylor, K.K.; Silva, J.M.N.; Sa, A.C.L.; Ni-Meister, W. A simulation analysis of the detectability of understory burns in miombo woodlands. *Remote Sens. Environ.* **2004**, *93*, 296–310. [[CrossRef](#)]
96. Genuer, R.; Poggi, J.-M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. [[CrossRef](#)]
97. Archer, K.J.; Kimes, R.V. Empirical characterization of random forest variable importance measures. *Comput. Stat. Data Anal.* **2008**, *52*, 2249–2260. [[CrossRef](#)]
98. Das, S. Filters, wrappers and a boosting-based hybrid for feature selection. In Proceedings of the 18th International Conference on Machine Learning (ICML), San Francisco, CA, USA, 28 June–1 July 2001; Citeseer: State College, PA, USA; pp. 74–81.
99. Liu, H.; Motoda, H. *Feature Selection for Knowledge Discovery and Data Mining*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; Volume 454.

