Unsupervised Burned Area Estimation through Satellite Tiles: A multimodal approach by means of image segmentation over remote sensing imagery

Alessandro Farasin[1,2], Giovanni Nini[2], Paolo Garza[1], and Claudio Rossi[2]

[1] Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy
{name.surname}@polito.it
[2] Fondazione LINKS, Via Pier Carlo Boggio, 61, 10138 Torino, Italy
{name.surname}@linksfoundation.com

Abstract. Climate change is increasing the number and the magnitude of wildfires, which become every year more severe. An accurate delineation of burned areas, which is often done through time consuming and inaccurate manual approaches, is of paramount importance to estimate the economic impact of such events. In this paper we introduce *Burned Area Estimation through satellite tiles* (BAE), an unsupervised algorithm that couples image processing techniques and an unsupervised neural network to automatically delineate the burned areas of wildfires from satellite imagery. We show its capabilities by performing an evaluation over past wildfires across European and non-European countries.

Keywords: burned area detection, self organizing maps, neural network, image segmentation, Copernicus EMS

1 Introduction

In the last years, we witnessed an increasing impact of natural hazards, such as wildfires and extreme weather events. In particular, the wildfire risk has increased also due to climate change, which is driving up temperatures, increasing the dry season. This trend poses several challenges with respect to the emergency management cycle, including the impact estimation of the event both in terms of human and economic losses. For wildfires, the economic impact is estimated starting from the delineation of the burned areas, which can be measured in several ways after the fire is completely extinguished. Usually, the burned area assessment is performed by domain experts working for public authorities, who adopt either an in-field approach, or manually draw an approximate polygon using satellite observations. Such manual annotations are inserted into local or national GIS and used for the impact assessment. This process is highly time consuming and error prone, negatively affecting the accuracy of the burned area definition, and consequently the impact estimation reliability. In the past, accurate supervised classification algorithms have also been proposed to address the burned

area delineation task (e.g., [6]). However, those techniques need large amount of precisely labeled/annotated images to train the classification algorithms. Since labeled data are not available for all areas of the world, which are characterized for instance by different vegetation patterns, supervised classifiers cannot be applied on new areas without an expensive initial manual labeling of several satellite images. To overcome the limitations of the above-mentioned manual and supervised approaches, we propose the Burned Area Estimation through satellite tiles (BAE) approach. BAE is an unsupervised automatic burned area delineation technique based on the coupled use of satellite imagery, image processing and neural networks. It exploits the availability of free high-resolution imagery provided by the optical satellites of the Sentinel-2 mission (Level 1C), which has been developed by ESA as part of the Copernicus Programme, to perform terrestrial observations. We performed a preliminary set of experiments to evaluate BAE, comparing its results with the official manual burned area annotations available in the Copernicus Emergency Management Service (EMS). Our preliminary experimental results show the ability of BAE to identify burned areas, excluding not affected land such as roads and rivers, which should not be considered in the impact assessment of wildfires. BAE is less accurate than supervised algorithms. However, it is independent of the area under observation and can be applied also to new areas for which labeled data are not available.
The rest of the paper is organized as follows. Section 2 introduces the related work, while Section 3 describes BAE. Section 4 reports the preliminary experimental results and Section 5 draws conclusions and presents possible future works.

2 Related Work

Different techniques have been proposed and exploited for addressing the burned area assessment problem [6, 11, 12, 3, 2, 1, 5, 8–10]. They can be classified in two main categories: (i) active fire mapping [11, 12, 3, 2] and (ii) post-wildfire mapping [6, 1, 5, 8–10]. Active fire techniques use the extension of active fires as a proxy for burned area estimation. However, those approaches overestimate the actual burned areas including also sub-areas that are covered by smoke but are not burned. On the other hand, the same approaches underestimate the burned area extension because the estimation is based on a subset of satellite images collected during the time period of the wildfires, the number of which could be limited either due to the satellite revisit time or to cloud obstruction. Post-wildfire techniques are usually more effective. They are supervised approaches that aim at building classification models able to label each image pixel as burned or not. The main limitation of those approaches is given by the need of a large training set of images containing samples of burned and unburned pixels. The manual annotation of training images is usually done by domain experts and it is time consuming. Hence, a limited amount of annotations is available. Moreover, the quality of the manually labeled images can be low due to human approximations. The low quality of the training data set has a significant negative impact

on the accuracy of supervised classification models. Another limitation of supervised approaches lies in their worldwide applicability, because the training of an accurate model requires the labeled set to contain all possible conditions, e.g., land type, vegetation intensity, which is impractical. For this reason, the supervised approaches proposed so far are usually focused on specific subareas of the world (e.g., Mithal et al. proposed a supervised approach specifically designed for burned areas in tropical forests [6]). We aim to overcome the mentioned limitations with BAE, an unsupervised post-wildfire approach based on satellite imagery, neural networks and image processing. The proposed approach does not need labeled data and can be applied worldwide, independently of the land type.

3 The BAE Framework

BAE is triggered after a wildfire event and works under the assumption to have an initial geo-graphical bounding box of interest and the dates at which the fire was ignited and extinguished. First, BAE uses the data services of the Satellite data provider to fetch two cloud free Red Green Blue (R, G, B) images having cloud coverage $< 5\%$. The choice to use the RGB bands is due to their availability in any satellite optical sensor and their derivability from other bands (such as Sentinel-2 B04, B03 and B02 bands, respectively). The first image must be dated before the fire ignition, while the second one after the fire was extinguished. Once the two images have been downloaded, BAE (represented in Figure 1) is applied to generate a binary mask (1-bit per pixel image) of the burned area, where White (W) and Black (B) pixels represent burned and unburned areas, respectively.
The BAE algorithm works as follows. First, the two gathered images are preprocessed separately by the Normalization & HSV Preprocessing module, which performs a Z-Score Normalization and a lossless conversion from RGB to Hue Saturation lightness Value model (HSV). Then it applies a transformation keeping the same H and setting both S and V to a constant value. We chose to select their maximum value (S_{max}, V_{max}) because this allows increasing the distance between color values in the RGB domain and hence the values can be clustered more easily in the following steps. This step is key to make the color component comparable between the two images while removing the differences that can result from images taken at different conditions (e.g. time of the day). The HSV Preprocessing model outputs the isolated H component, which is sent to the Hue Difference Segmentation Strategy module, and the (H,S_{max}, V_{max}), which is the input of the Color-based segmentation strategy module. Both strategies are detailed in the next subsections.

3.1 Hue Difference Segmentation Strategy (HDSS)

This strategy is based on the assumption that in an area affected by a wildfire the greatest changes in terms of pure colors (H) between the before image and

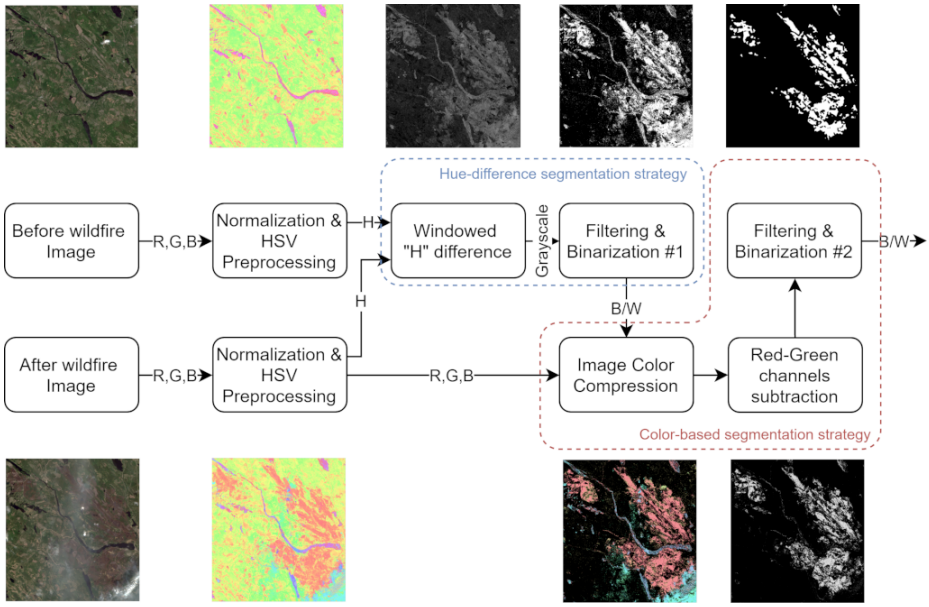


Fig. 1. Diagram of the BAE algorithm for segmenting burned areas. The rectangular boxes indicate the main algorithm steps, while the arrows describe the input/output types. The dotted-colored-boxes enclose the two different segmentation strategies.

the after image are due to the burned areas. However, not only wildfires produce significant changes of hue during a short period of time, but at this stage, this module aims to detect every area that has been subjected to a modification during the two times in which the before and after wildfire pictures were taken. Hence, the *Windowed "H" difference module* computes the difference between the "H" components of the two images. To reduce errors due to objects e.g., metallic surfaces, that change color when exposed to different kinds of sunlight, we take a 5x5 matrix of pixels in the before image and compute the minimum pixel wise difference with the pixel corresponding to the center of the matrix in the after image. We select 5x5 pixel matrix because this is the average size used in most image preprocessors. Let $HA_{i,j}$ be the "H" component of a pixel in position (i, j) in the pre-processed after-event image. Then, consider $W_{i,j}$ as a squared Window of odd size (w, w) centered in position (i, j) in the pre-processed before-event acquisition. This module generates a Hue Difference matrix HD having the same size of the input images, in which each pixel $HD_{i,j}$ represents the minimum distance between $HA_{i,j}$ and the pixel values p in $W_{i,j}$, which is computed as follows:

$$HD_{i,j} = min_{p \in W_{i,j}} angdist(HA_{i,j}, p), \tag{1}$$

where $angdist(x, y)$ is the angular distance between x and y. The angular distance is necessary because the H component is expressed in degrees, from 0 to

360. In HD, 0 values mean that there is no difference in that location with respect to the before-event situation, while positive values indicate the variation magnitude. Finally, the Filtering and Binarization #1 module applies a Gaussian Filter of dimensions (5,5), blurring the image and facilitating the computation for an automatic thresholding to binarize the image, performed by means of the standard Otsu's algorithm [7].

3.2 Color-based Segmentation Strategy (CSS)

This second strategy is based on the assumption that burned areas in the same image are characterized by similar colors. The first step of this strategy, named I*mage Color Compression*, works solely on the regions identified by the white color of the binary mask generated in the previous step. Accordingly to that mask, this module selects the colors of the HSV processed after wildfire image to reduce the color space and clusterize similar areas. Then, the second step, named *Red-Green Channels subtraction*, isolates the burned regions performing a subtraction between the Red and Green channels of the image. Finally, a second filtering and binarization step fine-tunes the segmentation. The mentioned steps are detailed in the following.
After converting the HSV after-event input image back to RGB, which we name Pre-processed Image (PI), the *Image Color Compression module* selects only the regions affected to a significant change, by selecting the colors in the regions resulted white in the binary image returned by the HDSS strategy. Then, it reduces the number of colors to "force" similar regions to be represented by the same RGB triple, i.e., we aim to "cluster" similar regions associating them to the same RGB triple. To accomplish this task, we adopt a Self-Organizing Map (SOM) [4], which is an unsupervised Artificial Neural Network (ANN) that maps the input image while preserving its neighborhood relations. The ANN neurons can be represented as a lattice of dimensions (n, m), in which each neuron is a multidimensional value (RGB in our case). First, the neurons are initialized in the multidimensional space. Then, the neuron that minimizes its average distance with the input image is selected as the BAE Matching Unit is (BMU), which will receive the maximum importance (weight in the ANN) and from which the training epochs (steps) will start. A neighborhood function modulates the weight update for the neurons of the whole network, keeping the weight updates inversely proportional to the distance from the BMU (see Figure 2). As explained before, the *HSV Preprocessing module* increases the pixels distance in the RGB space, facilitating the SOM training process in producing more representative neurons (see Figure 3).

The *Image Color Compression module* normalizes the input image by using the *min-max normalization*, which maps the RGB components from the range $[0, 255] \in \aleph$ to $[0, 1] \in \Re$ and feed that to the SOM, which should be carefully sized and initialized to be effective. We empirically set the network size to (3, 3), while we uniformly initialize the network weights in normalized RGB space. The network is trained for each image in an adaptive way with an increasing number

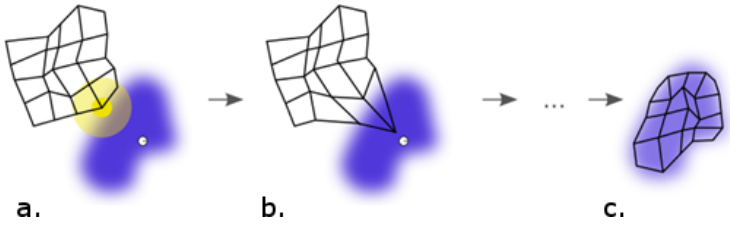
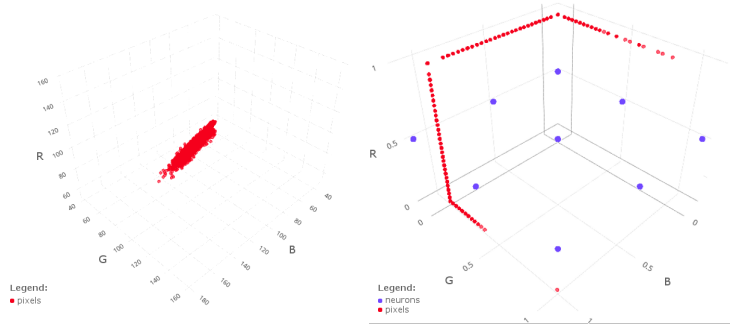


Fig. 2. Self-Organizing Map training process. The violet region represents the dataset, the grid represents the SOM of dimensions (5, 5) (having a neuron at each intersection). a) First step of the process, the BMU is the neuron surrounded by the yellow circle, its radius represents the neighborhood function which affects the other neurons weights update. b) Weights update result, after the previous epoch. c) The SOM after the training.



(a) (b)

Fig. 3. RGB representation for (a) the original and (b) preprocessed RGB after wildfire image. In (b), the initialized SOM neurons are represented in blue.

of epochs, until convergence. We show in Figure 3 the RGB representation of the after-wildfire image before and after the HSV Preprocessing, as well as the initialized SOM neurons. We select the Euclidean distance as the metric to perform the SOM training, after which each network neuron represents a color that summarizes a portion (pixel cluster) of the input space. We define TN as the set of Trained Neurons of the SOM and CC as the Color Compressed image that the SOM outputs. Every CC pixel $CC_{i,j}$ is assigned to the color of the neuron $n \in TN$ that minimizes the Euclidean distance from the corresponding $PI_{i,j}$ pixel:

$$CC_{i,j} = argmin_{n \in TN} ||PI_{i,j} - n||_2 \tag{2}$$

Therefore, in the white-colored regions of the HDSS output, CC is an RGB image having a reduced number of distinct colors equal to the number of the network neurons. While, in the remaining regions, it is colored in black.
The clusterization performed by the Image Color Compression module made similar colors closer to each other and, at the same time, it increased the dis-

tance respect to different colors. At this point, a module that allows highlighting common characteristics of burned regions is needed. In an unsupervised manner, this implies that no previous knowledge about the data can be exploited, but just a generic intuition is allowed. The idea behind the *Red-Green Channels subtraction* module is that, considering the hue of burned regions, they are prominent to red/violet colors and, at the same time, they present a near-to-zero level of green. Therefore, that module subtracts the green component from the red one. We do not consider the blue channel because, even if it is highly relevant in blue or light-blue areas like rivers or lakes, it is also relevant in violet regions, which can characterize burned areas. Finally, the *Filtering & Binarization #2* module is equivalent to the one adopted in the HDSS, with the addition in the end of a median filter, which remove possible noise generated by the binarization phase.

4 Preliminary Experimental Results

We performed a preliminary set of experiments to evaluate the quality of the burned areas identified by BAE with respect to the manual EMS annotations and the differences between the two approaches.

4.1 Evaluation Dataset

To carry out the experimental comparison between BAE and EMS manual annotations, the EMS annotations and the related satellite images were collected on the basis of a two-step process.
The first step is based on the retrieval of metadata information describing the historical summary of wildfire events from 2016 to 2018 thanks to the Emergency Management System service of the Copernicus Program (emergency.copernicus.eu). EMS provides several information for each single wildfire event, describing, through a manual annotation based on a polygon represented in EPSG:4326 coordinates, the area hit by the fire. This information was used for creating a wider bounding box, surrounding the burned area, obtained by doubling, respectively, the differences between the lower and upper latitudes and the leftmost and rightmost longitudes of the given polygon. The approach used for defining the bounding box allows including also burned areas that are erroneously outside of the manual annotation. Finally, the dates representing the start and the end of the wildfire event are also collected. The second step fetches the satellite images associated with the wildfires of interest, exploiting the previously collected event metadata. The number of analyzed wildfire events is 50, covering 10 countries representative of different land types. The processed satellite images cover areas of the following countries: Cyprus, Georgia, Greece, Italy, Macedonia, Morocco, Portugal, Spain, Sweden, and Tunisia. For each of the 50 available wildfire events, we applied BAE and then we compared the EMS manual annotated burned areas with those identified by BAE. We initially performed a subjective analysis by manually comparing the two annotations and then an objective one based on recall and precision by considering the EMS annotations as ground-truth.

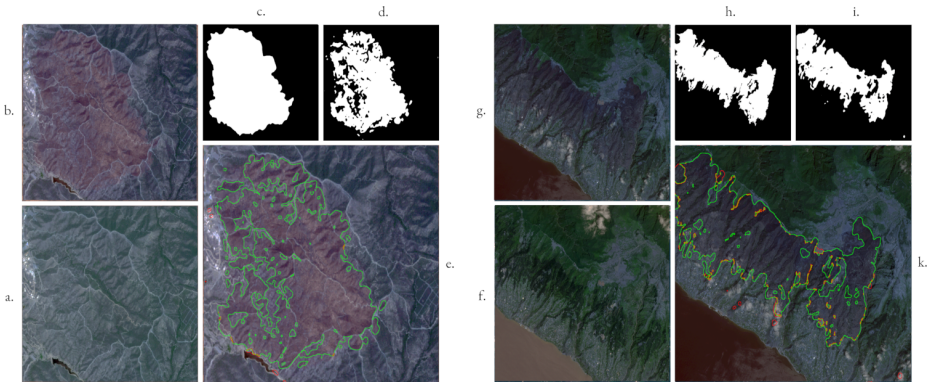


Fig. 4. Comparison between BAE and EMS annotations for two representative wildfires. (a,f) Before wildfire event. (b,g) After wildfire event. (c,h) EMS annotations: white pixels indicate burned areas. (d,i) BAE. (e,k) Burned areas identified by BAE: green areas represent the intersection with the EMS annotations, while the red ones represent areas identified by BAE, but excluded by EMS.

4.2 Comparison between BAE and EMS annotations

We report the images and annotations associated with two representative wildfire events to show the differences between the burned areas identified by BAE and the manual EMS annotations (see Figures 4.a-4.e for the first example and Figures 4.f-4.k for the second one). Let us consider the first example (see Figures 4.a-4.e). We can notice that the manual annotation (Figure 4.c) is one single completely filled polygon with a smoothed shape, whereas BAE is more precise and identifies a set of highly detailed irregular shapes representing several burned subareas (Figure 4.d). Specifically, the manual EMS-based annotation labels as burned areas also some inner spots of the burned area polygon, which are clearly unburned, whereas BAE correctly label those pixels as unburned. Moreover, BAE correctly labels as unburned pixels also some border regions, which are erroneously labeled as burned in EMS. Figures 4.f-4.k show a second representing example that highlights another difference between BAE automatic delineations and EMS annotations. In this specific case, BAE identifies also some small-burned areas that are not reported in the human generated annotation (e.g., a small white pixel in the bottom-right corner of Figure 4.k). To highlight differences between the two approaches, the intersection of the EMS annotation and the BAE annotation is marked as green, while burned pixels annotated by BAE that are not present in the manual one are marked as red (see Figure 4.e for the first wildfire event example and Figure 4.k for the second one). For the 50 analyzed wildfires, we computed also a set of objective statistics to evaluate similarities and differences between the burned areas identified by BAE and the manually labeled ones available in EMS. Specifically, we considered the EMS annotations as our ground-truth and we computed the standard recall, precision and F-score measures for the annotations generated by BAE. Given an arbitrary

wildfire event, and the associated after wildfire image, we consider two sets of pixels: B_{BAE} and B_{EMS}, which are the sets of pixels labeled as burned by BAE and EMS, respectively. Based on those two sets, recall, precision and F-score of the burned class have been computed. The average precision computed over the considered 50 acquisitions is 81%. Hence, on the average, 81% of the pixels labeled as burned by BAE are also considered burned areas by EMS. In some cases, BAE correctly identifies some small subareas that are not included in the manual annotations, like we discussed above (i.e., the erroneous information is in the EMS annotation in that case). In some other cases, BAE erroneously labeled as burned some pixels that have not been affected by a wildfire. The error is related to a significant difference between the before and after images that is not related to a wildfire. In terms of recall, BAE achieves an average recall value equal to 66%. Hence, only 66% of the pixels manually annotated in EMS as burned areas are labeled as burned also by BAE. Also in this case, in some cases the error is given by not precise manual annotations (for instance because small unburned subareas of a large burned region are not properly annotated as unburned in EMS). In some other cases, the error is due to BAE. Finally, the average F-score is 70%. Moreover, we evaluated the contribution introduced by the color compression performed by the SOM in the Image Color Compression module, compared to a simpler version in which the Red-Green subtraction module operates directly on the colors identified by the output of HDSS. As a result, without the color compression, the average F-Score lowers to 67%, loosing 5-6% on average recall. To assess the impact of the two approaches on the estimation of the burned areas in km^2, we finally computed, for the considered 50 wildfire events, the total number of pixels labeled as burned by BAE and EMS, respectively. The number of pixels labeled as burned by BAE is 4'193'548, while EMS labels 5'351'961 pixels as burned. Since each pixel corresponds approximately to 100 m^2, the estimation of the amount of burned areas based on the manual EMS annotations is approximately 535 km^2, where as the one obtained by means of BAE is only 419 km^2 (i.e., the estimation based on the EMS annotations is 1.3 times larger than the BAEs one). Moreover, the area labeled as burned by both solutions is only 326 km^2. These results highlight the significant difference between the two approaches and hence the correct evaluation of wildfire impacts.

5 Conclusions and Future Work

In this paper, an unsupervised burned area detection framework, based on the combination of a *hue difference segmentation strategy* and a *color-based segmentation strategy*, has been proposed. The proposed approach is less accurate than manual annotations. However, it allows automating the burned area identification task without the human intervention. As future work, we plan to exploit the additional bands provided by Sentinel-2 to further improve the quality of BAE. Moreover, the exploitation of supervised deep neural network techniques, without the need of before-event images, represents a natural continuation of

this work, exploiting the annotations achieved through the BAE framework as training data.

References

1. 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 Sensing **6**(12), 12360–12380 (2014)
2. Fraser, R., Li, Z., Cihlar, J.: Hotspot and ndvi differencing synergy (hands): A new technique for burned area mapping over boreal forest. Remote Sensing of Environment **74**(3), 362–376 (2000)
3. 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 Sensing of Environment **113**(2), 408–420 (2009)
4. Haykin, S.S., Haykin, S.S., Haykin, S.S., Elektroingenieur, K., Haykin, S.S.: Neural networks and learning machines, vol. 3. Pearson education Upper Saddle River (2009)
5. Loboda, T., Csiszar, I.: Reconstruction of fire spread within wildland fire events in northern eurasia from the modis active fire product. Global and Planetary Change **56**(3-4), 258–273 (2007)
6. Mithal, V., Nayak, G., Khandelwal, A., Kumar, V., Nemani, R., Oza, N.: Mapping burned areas in tropical forests using a novel machine learning framework. Remote Sensing **10**(1), 69 (2018)
7. Otsu, N.: A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics **9**(1), 62–66 (1979)
8. Pu, R., Gong, P.: Determination of burnt scars using logistic regression and neural network techniques from a single post-fire landsat 7 etm+ image. Photogrammetric Engineering & Remote Sensing **70**(7), 841–850 (2004)
9. Pu, R., Gong, P., Li, Z., Scarborough, J.: A dynamic algorithm for wildfire mapping with noaa/avhrr data. International Journal of Wildland Fire **13**(3), 275–285 (2004)
10. Roy, D.P.: Multi-temporal active-fire based burn scar detection algorithm. International Journal of Remote Sensing **20**(5), 1031–1038 (1999)
11. Schultz, M.: On the use of atsr fire count data to estimate the seasonal and interannual variability of vegetation fire emissions. Atmospheric Chemistry and Physics **2**(5), 387–395 (2002)
12. Sukhinin, A.I., French, N.H., Kasischke, E.S., Hewson, J.H., Soja, A.J., Csiszar, I.A., Hyer, E.J., Loboda, T., Conrad, S.G., Romasko, V.I., et al.: Avhrr-based mapping of fires in russia: New products for fire management and carbon cycle studies. Remote Sensing of Environment **93**(4), 546–564 (2004)