Global estimation of burned area using MODIS active fire observations

L. Giglio, G. R. van Der Werf, J. T. Randerson, G. J. Collatz, P. Kasibhatla

Atmos. Chem. Phys., 6, 957–974, 2006
www.atmos-chem-phys.net/6/957/2006/

Atmospheric Chemistry and Physics

Global estimation of burned area using MODIS active fire observations

L. Giglio[1], G. R. van der Werf[2], J. T. Randerson[3], G. J. Collatz[4], and P. Kasibhatla[5]

[1]Science Systems and Applications, Inc., NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[2]Department of Hydrology and Geo-Environmental Sciences, Vrije Universiteit, Amsterdam, The Netherlands
[3]Department of Earth System Science, University of California, Irvine, California, USA
[4]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[5]Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina, USA

Received: 2 September 2005 – Published in Atmos. Chem. Phys. Discuss.: 1 November 2005
Revised: 13 January 2006 – Accepted: 6 February 2006 – Published: 28 March 2006

Abstract. We present a method for estimating monthly burned area globally at 1° spatial resolution using Terra MODIS data and ancillary vegetation cover information. Using regression trees constructed for 14 different global regions, MODIS active fire observations were calibrated to burned area estimates derived from 500-m MODIS imagery based on the assumption that burned area is proportional to counts of fire pixels. Unlike earlier methods, we allow the constant of proportionality to vary as a function of tree and herbaceous vegetation cover, and the mean size of monthly cumulative fire-pixel clusters. In areas undergoing active deforestation, we implemented a subsequent correction based on tree cover information and a simple measure of fire persistence. Regions showing good agreement between predicted and observed burned area included Boreal Asia, Central Asia, Europe, and Temperate North America, where the estimates produced by the regression trees were relatively accurate and precise. Poorest agreement was found for southern-hemisphere South America, where predicted values of burned area are both inaccurate and imprecise; this is most likely a consequence of multiple factors that include extremely persistent cloud cover, and lower quality of the 500-m burned area maps used for calibration. Application of our approach to the nine remaining regions yielded comparatively accurate, but less precise, estimates of monthly burned area. We applied the regional regression trees to the entire archive of Terra MODIS fire data to produce a monthly global burned area data set spanning late 2000 through mid-2005. Annual totals derived from this approach showed good agreement with independent annual estimates available for nine Canadian provinces, the United States, and Russia. With our data set we estimate the global annual burned area for the years 2001–2004 to vary between 2.97 million and 3.74 million km^2, with the maximum occurring in 2001. These coarse-resolution burned area estimates may serve as a useful interim product until long-term burned area data sets from multiple sensors and retrieval approaches become available.

Correspondence to: L. Giglio
(giglio@hades.gsfc.nasa.gov)

1 Introduction

Research over the past 25 years has led to increased recognition of the important role biomass burning plays in the global carbon cycle and the production of trace gas and aerosol emissions. Consequently, Earth-system modeling efforts now often include fire-related information. In particular, there is a strong need for spatially and temporally explicit estimates of the quantity of biomass consumed through combustion (Scholes et al., 1996). Typically such estimates are based on a simple relationship of the form (e.g., Seiler and Crutzen, 1980; Hao et al., 1990; Pereira et al., 1999)

$$M = ABc, \tag{1}$$

where M is the mass of vegetation combusted within a given time interval, A is the area burned during the same time interval, B is the biomass density, and c is a factor describing the completeness of combustion. Although all of the terms appearing on the right hand side of Eq. (1) are highly variable, burned area is particularly difficult to estimate because of the potentially high spatial and interannual variability in this quantity at continental to global scales. It is therefore especially important that accurate, spatially explicit, multi-year estimates of burned area are available when relying on a relationship having the form of Eq. (1). At present, however, there is a dearth of such data. While a

Published by Copernicus GmbH on behalf of the European Geosciences Union.

number of satellite-based global burned area products are currently under development, specifically GLOBSCAR (Simon et al., 2004), GBA2000 (Tansey et al., 2004), and the MODIS burned area product (Justice et al., 2002; Roy et al., 2002), none are yet available on a multi-year basis.

Unlike burned area data, long-term observations of active fires made with spaceborne sensors are readily available. Representative multi-year examples include the Along-Track Scanning Radiometer (ATSR) nighttime fire product (Arino and Rosaz, 1999), the Visible and Infrared Scanner (VIRS) monthly fire product (Giglio et al., 2003a), the Moderate Resolution Imaging Spectroradiometer (MODIS) global fire product (Justice et al., 2002), and the Geostationary Operational Environmental Satellite (GOES) Wildfire Automated Biomass Burning Algorithm (WF_ABBA) fire product (Prins et al., 1998). At their most basic level, active fire products contain information about the location and timing of fires that are burning at the time of the satellite overpass, usually in the form of swath-based fire masks or as lists of fire pixel locations and dates. These observations are in turn often summarized at coarse spatial resolutions (e.g., $0.5^\circ \times 0.5^\circ$) over daily or monthly time periods, yielding data products containing gridded counts of active fire pixels. Although these "fire count" products capture many aspects of the spatial distribution and seasonality of burning, it is difficult to relate them to actual area burned due to inadequate temporal sampling, variability in fuel conditions and cloud cover, differences in fire behavior, and issues related to spatial resolution (Scholes et al., 1996; Eva and Lambin, 1998; Kasischke et al., 2003).

Despite these difficulties, the lack of long-term, spatially-explicit global burned area data has meant that active fire observations must often be used as a proxy for area burned (e.g., Setzer and Pereira, 1991; Scholes et al., 1996; Stroppiana et al., 2000; Potter et al., 2001; van der Werf et al., 2003, 2004; Langmann and Heil, 2004). Perhaps the most common approach has been to assume that the area burned is proportional to simple counts of fire pixels, i.e.

$$A(i,t) = \alpha N_{\mathrm{f}}(i,t), \tag{2}$$

where A is the area burned within a particular spatial region labeled by the index i – typically a grid cell – during a fixed time period labeled by the index t, N_{f} is the number of fire pixels observed within the same region during the same time period, and α is a constant representing the effective burned area per fire pixel.

The reported accuracies of the burned area estimates obtained with Eq. (2) vary greatly and are dependent upon, among other things, the spatial scale at which the relationship is applied. Eva and Lambin (1998) found almost no correlation between AVHRR fire counts and burned area in the Central African Republic at a spatial resolution of 15 km over a time interval of about one month. Randriambelo et al. (1998), however, report a good qualitative agreement between one year of monthly AVHRR fire counts and ground-based monthly burned area estimates for a study region in Madagascar. Pereira et al. (1999) report a poor linear correlation (r=0.44) between daytime AVHRR fire counts and burned area estimates in a 20° by 10° region encompassing the Central African Republic over a 25-day time period. Kasischke et al. (2003) examined the relationship between ATSR fire counts and area burned in Alaska and Canada from 1997 to 2002, and in Russia during 1998. They reported significant linear correlations between fire counts and burned area for Canada and Russia, but in the former region found that the slope (i.e., the effective area burned per fire pixel) for different years varied by up to a factor of about two. The authors caution against scaling fire counts to area burned since rates of fire detection, cloud obscuration, and fire spread are not constant across years.

Variations of Eq. (2) in which α assumes some spatial dependence have also been explored. Scholes et al. (1996) were able to relate the area burned in southern Africa to monthly 0.5° gridded AVHRR fire counts using ancillary Normalized Difference Vegetation Index (NDVI) data such that $\alpha(i) = f[\mathrm{NDVI}(i)]$, where f is a linearly decreasing function of the mean annual NDVI in grid cell i. In other words, increasing greenness reduces the effective burned area per fire pixel. Van der Werf et al. (2003) related burned area to VIRS active fire counts using fractional tree cover at a spatial resolution of 1° such that $\alpha(i) = f[T(i)]$, where f is a linearly decreasing function of the mean fractional tree cover T in grid cell i. Here, increasing tree cover slows the fire spread rate and reduces the effective burned area per fire pixel. The two approaches are closely related since NDVI and tree cover are positively correlated.

Active fire observations have also been used to spatially and temporally allocate climatological inventories of combusted biomass and pyrogenic trace gas emissions (Schultz, 2002; Duncan et al., 2003; Generoso et al., 2003; Heald et al., 2003; Streets et al., 2003). These methods are fundamentally related to Eq. (2) in that they assume the quantity of interest is proportional to counts of fire pixels. While our interest here is confined to burned area, much of the subsequent discussion is applicable to allocation-based approaches as well.

In this paper, we present a method for calibrating active fire observations made with the Terra MODIS sensor to produce global, coarse resolution estimates of burned area on a monthly basis. Our approach draws upon two types of information: the sensitivity of α to fractional tree and herbaceous cover (extending the approach used by van der Werf et al., 2003), and the sensitivity of α to fire-pixel cluster size. These components were combined using regression trees that were applied to large geographic regions. In recognizing that production of accurate burned area maps suitable for calibration is problematic in closed canopy tropical forest, particularly in areas of active deforestation, we implement a subsequent refinement in which a correction is applied to the burned area predicted with the regression trees using tree cover data and a simple measure of fire persistence.

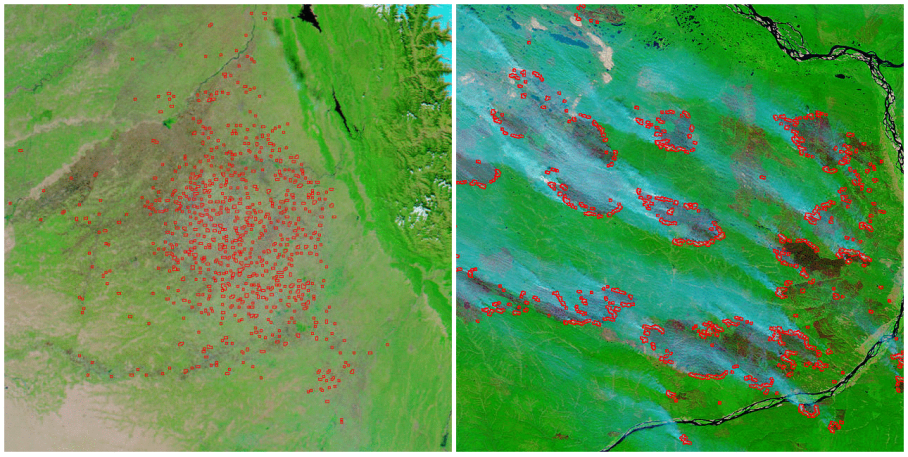

Fig. 1. Aqua MODIS 500-m false color imagery of northern India (left) on 23 October 2004 (08:20 UTC) and Yakutsk, Russia (right) on 19 August 2002 (03:00 UTC). Outlines of 1-km active fire pixels are shown in red. With this band combination (2.1 μm, near-infrared, red) dense vegetation appears green, heavy smoke appears light blue, burn scars appear dark brown, water appears black, and non-cirrus clouds appear white. The scale (approximately 320×320 km) is identical in both images. Note how larger Yakutsk burn scars are accompanied by large clusters of adjacent fire pixels, while the small (but numerous) agricultural burns in India are characterized by much smaller clusters of fire pixels and no visible burn scars. Images were produced within the MODIS Rapid Response System and appear courtesy of Jacques Descloitres.

The uncertainties associated with any calibration approach are likely to be comparatively large given the sampling issues mentioned above, but for some applications may still be tolerable. Global models of the terrestrial carbon cycle, for example, have only recently begun to include explicit treatment of fire as a disturbance factor (e.g., van der Werf et al., 2003). We expressly do not claim that an active-fire based method can provide a universal substitute for burned area maps generated via direct observation of burn scars. Rather, in agreement with Schultz (2002), we suggest that statistical coarse-resolution burned area estimates derived from MODIS active fire observations can serve as a useful interim product until long-term burned area data sets become available. Moreover, it may be possible to use an active-fire calibration approach with functional sensors pre-dating MODIS, offering the possibility of generating even longer-term global burned area data sets.

2 Data

2.1 Active fire data

We used the Collection 4, version 4 Terra MODIS monthly Climate Modeling Grid (CMG) fire products at 0.5° spatial resolution ("MOD14CMH"), from January 2001 through December 2004. The gridded monthly overpass-corrected fire pixel counts were summed to a 1° working spatial resolution for this study. The CMG product also contains the mean percent tree cover (T_f), percent herbaceous vegetation cover (H_f), and percent bare ground (B_f) from the global MODIS Vegetation Continuous Fields (VCF) products (Hansen et al., 2003) for all fire pixels within each grid cell; we averaged these to 1° spatial resolution as well. (We use the subscript "f" as a reminder that T_f, H_f, and B_f are averages for fire pixels only, as opposed to averages over the entire land surface encompassed by the grid cell.)

Using the locations of individual Collection 4 MODIS fire pixels at the nominal 1-km MODIS resolution (available separately), we linked adjacent fire pixels within each 1° grid cell into clusters on a monthly basis. For each grid cell we then computed the monthly mean fire-pixel cluster size, which we denote as C_f. We hypothesized that cluster-related information might improve the estimation of burned area based on the empirical observation that larger clusters of MODIS fire pixels tend to be associated with larger burn scars (Fig. 1).

2.2 Burned area data

Burned area maps were produced using a prototype algorithm that uses the 500-m MODIS atmospherically-corrected Level 2G surface reflectance product (Vermote et al., 2002), the MODIS Level 3 daily active fire products (Justice et al., 2002), and the MODIS Level 3 96-day Land Cover Product (Friedl et al., 2002). The algorithm, which is described in the Appendix, identifies the date of burn, to the nearest day, for pixels within individual MODIS Level 3 tiles (Wolfe et al., 1998) at 500-m spatial resolution. Since these burn scar masks were to serve as truth for calibration of active fire observations, we visually inspected each to ensure that no obvious omission or commission errors were present. Often

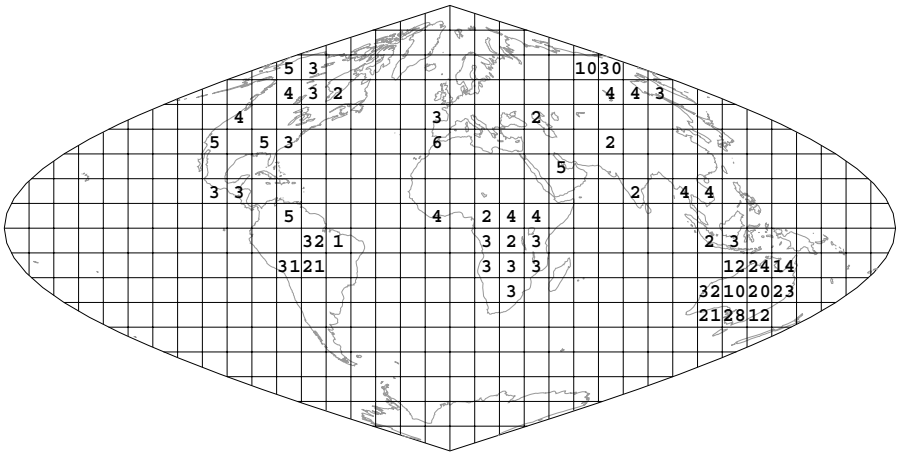

Fig. 2. Locations of MODIS calibration tiles used in this study. Numbers in each $10° \times 10°$ tile indicate the number of months for which 500-m burned area masks were produced for that tile.

this required appealing to higher resolution 250-m MODIS imagery to verify the existence of smoke plumes and help resolve the boundaries of ambiguous burn scars. Manual corrections were required in approximately five tiles, usually to add a burn scar that was undetected due to persistent cloud cover. At present, validation of our 500-m burned area maps has been limited to Russia through comparison with maps generated manually from high resolution Landsat imagery (Loboda and Csiszar, 2004). Proper global validation would require that a similar procedure be applied to representative sites over the entire globe. This is a very substantial undertaking that has not yet been completed for any burned area product.

Selected calendar months were processed for selected MODIS tiles, yielding a total of 446 "tile-months" of burned area estimates between January 2001 and December 2004 (Fig. 2). Tile locations were selected to provide a good sampling of worldwide fire activity over multiple fire seasons, although erratic data availability ultimately produced an uneven temporal sampling of the different tiles. The resulting burned area maps were aggregated to 1° spatial resolution and monthly temporal resolution.

While we believe that commission and omission errors in our 500-m burned area maps are generally negligible compared to the statistical variability inherent in modeling the relationship between burned area and active fire pixels with Eq. (2), we recognize that the quality of these maps is substantially lower in the closed canopy forests of South America and Equatorial Asia. A combination of three factors make mapping of burned area problematic in this biome. First, surface burns are at least partially obscured by the tree canopy, which can leave an insufficient post-burn, top-of-atmosphere radiometric signal with which to detect the burn. Second, substantial spectral overlap can occur between cleared (but unburned) forest patches, and patches that have been cleared and subsequently burned. Finally, persistent cloud cover ($\sim$1 month and longer) is common in rainforest, and this can lead to significant errors of omission, particularly following vegetation regrowth. This issue will be addressed further in Sect. 3.3.

3 Method

3.1 Preliminary analysis

As part of a preliminary analysis we examined the relationship between monthly corrected Terra fire pixel counts and area burned within 14 different regions (Fig. 3) using the model in Eq. (2). Results obtained from least squares fits to this model are summarized in Table 2. There is clearly strong regional variation in the effective area per fire pixel (α), from a minimum of 0.29 km^2/pixel in southern-hemisphere (SH) South America to a maximum of 6.6 km^2/pixel in Central Asia, which is a factor of more than 20. We note that, with the exception of the SH South America region, the correlation coefficients we obtained are substantially higher than those reported by Boschetti et al. (2004) between ATSR fire counts and the GBA2000 and GLOBSCAR burned area data sets for the year 2000. There are at least five possible reasons for our higher correlation. First, our 1° grid cells are larger than the hexagonal grid cells used by Boschetti et al. (2004) by about a factor of four at the Equator, and by a factor of two at boreal latitudes. The correlation between many spatial quantities tends to improve over larger areas (Eva and Lambin, 1998). Second, with the exception of Europe, the six geographic regions defined by Boschetti et al. (2004) were much larger than the 14 geographic regions used in our study. Our results show that α can vary by as much as a factor of nearly seven within these larger regions. Third, the larger MODIS swath yields a higher temporal sampling

Table 1. Regions used within this study. Abbreviations refer to those used in Fig. 3.

Abbrev.	Short Name	Comments
BONA	Boreal North America	Alaska and Canada.
TENA	Temperate North America	Conterminous United States.
CEAM	Central America	Mexico and Central America.
NHSA	Northern Hemisphere South America	Division with SHSA is at the Equator.
SHSA	Southern Hemisphere South America	Division with NHSA is at the Equator.
EURO	Europe	Includes the Baltic States but excluding White Russia and the Ukraine.
MIDE	Middle East	Africa north of the Tropic of Cancer, and the Middle East plus Afghanistan.
NHAF	Northern Hemisphere Africa	Africa between the Tropic of Cancer and the Equator.
SHAF	Southern Hemisphere Africa	
BOAS	Boreal Asia	Russia, excluding area south of 55° N between the Ukraine and Kazakhstan.
CEAS	Central Asia	Mongolia, China, Japan, and former USSR except Russia.
SEAS	Southeast Asia	Asia east of Afghanistan and south of China.
EQAS	Equatorial Asia	Malaysia, Indonesia, and Papua New Guinea.
AUST	Australia	Includes New Zealand.

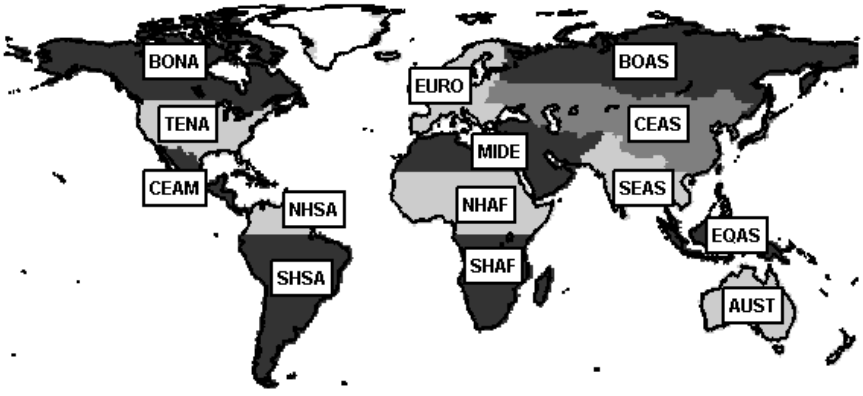

Fig. 3. Map of the 14 regions used in this study. Abbreviations are explained in Table 1.

rate, making it more likely that MODIS will "fill in" large burned areas with active fire pixels, and leading to fewer small burned areas for which no active fire pixels were detected. (We will return to the issue of temporal sampling in Sect. 7.) Fourth, the "fill-in" effect might become more pronounced in those regions having a strong diurnal fire cycle since fewer fires are likely to be burning at the time of the nighttime ATSR overpass. Finally, smaller burns present in the 500-m MODIS burned area maps might not be identified in the 1-km GBA2000 and GLOBSCAR data sets. This last factor contributes because, for pixels of a given size, the minimum detectable size of an actively burning fire is much smaller than the minimum detectable size of a burn scar (by a factor of $\sim$1000). Mapping burn scars with larger pixels will therefore yield more cases in which small clusters of active-fire pixels are not accompanied by an observable burn scar, and will therefore reduce the correlation between the two variables (cf. Fig. 1, left panel).

We repeated the above analysis with fire pixel counts having an additional correction for cloud cover (data layers with and without this correction are present in the MODIS CMG fire products). The resulting correlation coefficients were almost uniformly lower, most likely because the cloud correction relies on assumptions that are frequently not met and consequently has a tendency to overcorrect. We performed the remainder of our investigation, therefore, with overpass-corrected fire pixel counts lacking the additional cloud correction.

We next examined the effect of tree cover on α. We partitioned the observations for each region into 20% tree-cover intervals and fitted Eq. (2) separately to each of the resulting subsets. Results for five regions are shown in Fig. 4. We found that in savanna regions α decreased with increasing tree cover, although the slope of the relationship varied substantially between different savanna regions. A similar analysis revealed a comparable link between α and herbaceous cover, but with α increasing with increased herbaceous cover.

Table 2. Correlation (r) between predicted and observed burned area within each region for linear regression (Eq. (2)) and regression tree approaches. The slope of the linear regression (α) and the total number of non-zero observations (N) are also shown. An observation consists of the corrected fire pixel counts, 500-m burned area, mean VCF fraction for all fire pixels, and mean fire-pixel cluster size within a single 1° grid cell for a specific month. Observations having zero burned area and zero fire pixels ("zero-zero" observations) were not included in the analysis and are not reflected in the tabulated values of N. All correlations are highly significant with a probability $p \ll 0.001$.

	Linear (Eq. (2))			Tree
Region	α (km^2/pixel)	r	N	r
Boreal North America	1.4	0.69	1018	0.85
Temperate North America	0.84	0.94	982	0.98
Central America	0.43	0.73	301	0.85
NH South America	1.0	0.78	352	0.85
SH South America	0.29	0.35	4034	0.56
Europe	3.1	0.91	225	0.95
Middle East	0.40	0.34	215	0.78
NH Africa	5.2	0.86	910	0.90
SH Africa	2.9	0.60	1670	0.73
Boreal Asia	1.3	0.90	2104	0.94
Central Asia	6.6	0.85	282	0.92
South Asia	2.9	0.75	531	0.83
Equatorial Asia	0.49	0.71	192	0.78
Australia	3.4	0.82	5563	0.89

This is not surprising given that, in most fire susceptible areas, woody-herbaceous gradients (rather than, say, woody-bare gradients) are more often the norm. The variation in α with respect to bare cover was generally much weaker except in Australia, where fires are common along gradients of bare and herbaceous cover. In tropical forests (e.g., South America), there was no significant relationship between α and tree cover, and in boreal forests α slightly increased with increasing tree cover.

We also examined how α varied with respect to mean fire-pixel cluster size. We partitioned the observations for each region into different ranges of mean cluster size and fitted Eq. (2) to each of the resulting subsets. We found that, in general, the effective burned area per fire pixel increased very rapidly as cluster size increased (Fig. 5).

3.2 Regression tree approach

The analysis in the previous section shows that vegetation fraction (tree [T_f], herbaceous [H_f], bare [B_f]) and fire cluster size (C_f) are important predictive variables that should potentially appear in an empirical relationship linking active fire counts to area burned. We may write such a relationship very generally as

$$A(i,t) = g[T_f(i,t), H_f(i,t), B_f(i,t), C_f(i,t), N_f(i,t)] \quad (3)$$

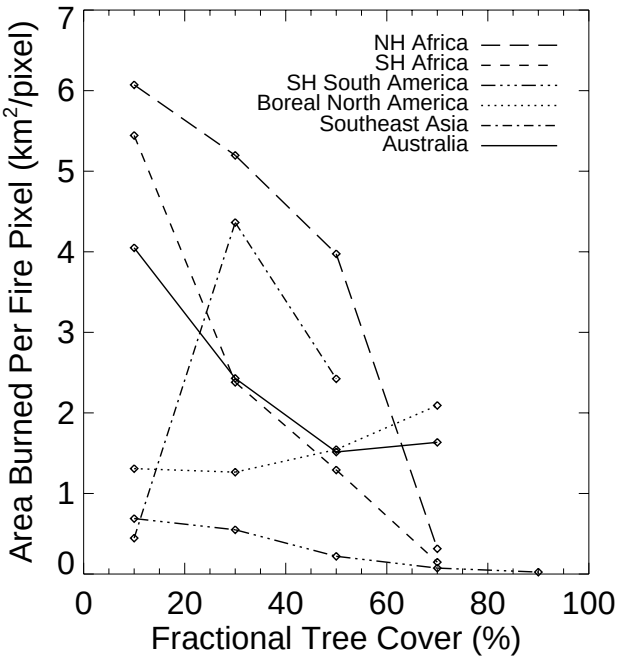

Fig. 4. Effective burned area per Terra MODIS fire pixel (α) as a function of mean percent tree cover for six of the 14 regions considered in this study.

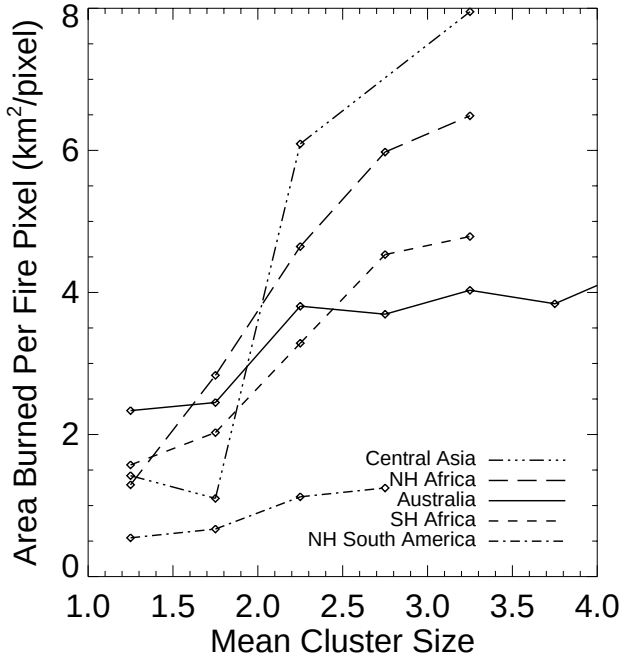

Fig. 5. Effective burned area per Terra MODIS fire pixel (α) as a function of mean fire-pixel cluster size for five of the 14 regions considered in this study.

where g is an unknown function. It is not obvious, however, what particular functional form one should assume for

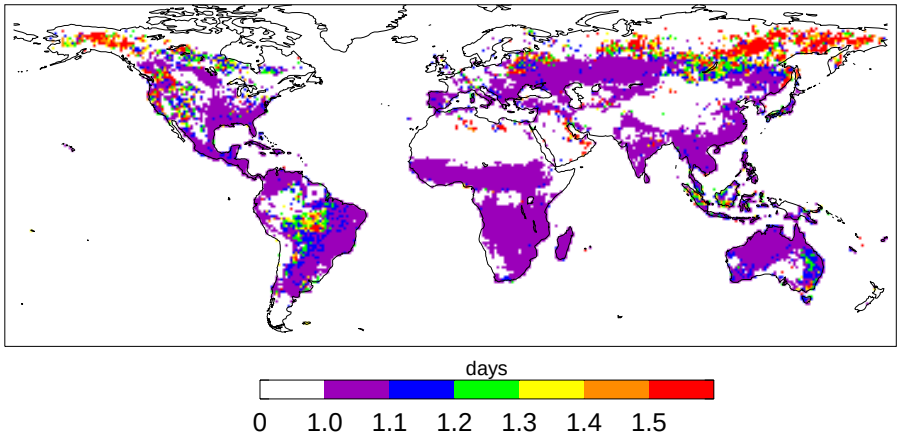

Fig. 6. 2001–2004 mean monthly fire persistence computed from Terra MODIS active fire observations.

g that will be optimal in every region. Based on a separate exploratory analysis, we believe that a globally optimal function is likely to require an unreasonably large number of free parameters. We therefore pursued the conceptually simpler approach of expressing the relationship in Eq. (3) as a *regression tree* for each region.

A regression tree is an alternative model for expressing a relationship between a continuous dependent variable y and one or more predictive (or explanatory) variables x_i (Breiman et al., 1984). The tree per se consists of a set of rules of the form "if $x_1<1$ and $x_2<2$ then $y=3$" which supply an appropriate value for y over the range of the x_i. These rules are constructed by partitioning (or splitting) observations along the x_i into two subsets in such a way as to maximize the reduction in an error metric (or "deviance"). Following the split, the homogeneity of the resulting pair of subsets is increased. This procedure is applied recursively to each subset until certain stopping criteria are met (typically the number of remaining observations becomes too small, or the reduction in deviance becomes insignificant). The resulting binary tree consists of splits (e.g., "if $x_1<1$"), and leaves (or terminal nodes) in which the dependent variable is assigned a value. Following tree construction, pruning is usually applied to eliminate overfitting that would otherwise degrade the predictive ability of the tree. During this process, terminal nodes having little predictive robustness are eliminated through the use of a cost-complexity function (Breiman et al., 1984).

For this study we used a more flexible form of regression tree which models the dependent variable using a linear regression in each terminal node (Breiman and Meisel, 1976). Trees built in this manner are usually smaller and are also often easier to interpret. The particular linear model we used was simply that in Eq. (2). To help ensure the resulting fit was robust, we required a minimum of 30 observations within each terminal node. (By "observation" we are referring to the corrected fire pixel counts, 500-m burned area, mean fire-pixel VCF fractions (tree, herbaceous, bare), and mean fire-pixel cluster size within a single 1° grid cell for a single month.)

During tree construction, we permitted splitting on all five predictive variables appearing in Eq. (3). These variables are clearly not independent given their constraints ($T_f+H_f+B_f=100\%$, $C_f \leq N_f$), and that, within the tropics, larger fire clusters tend to occur in regions having higher herbaceous cover. This multi-collinearity will have no impact on the predictive ability of the regression trees that we derive, but it does mean that the final choice of splitting variable will be more or less arbitrary in the event two such variables yield comparable reductions in deviance following a trial split. When interpreting the final trees, therefore, one should not attach too much significance to the fact that, say, tree cover was selected as the splitting variable rather than herbaceous cover.

3.3 Tropical closed-canopy forest correction

As mentioned in Sect. 2.2, accurate mapping of burned areas within tropical closed-canopy forest is extremely challenging. In brief, obscuration of the surface by persistent cloud cover and the tree canopy can lead to significant errors of omission. This problem is not unique to the MODIS instrument. A further complication occurs in tropical areas undergoing deforestation: where fire is used in the deforestation process (e.g., South America and Equatorial Asia), burning is usually preceded by mechanical clearing and aggregation of the resulting slash. Consequently, despite the fact that a relatively large area of forest has been cleared and burned, the spatial extent of the burn scar per se is much smaller than the area cleared. Satellite-based maps of burn scars under these conditions are therefore likely to systematically underestimate the effective area burned and fuel consumed. To help rectify this in our burned area product, we attempt to use information about fire persistence and tree cover to specify

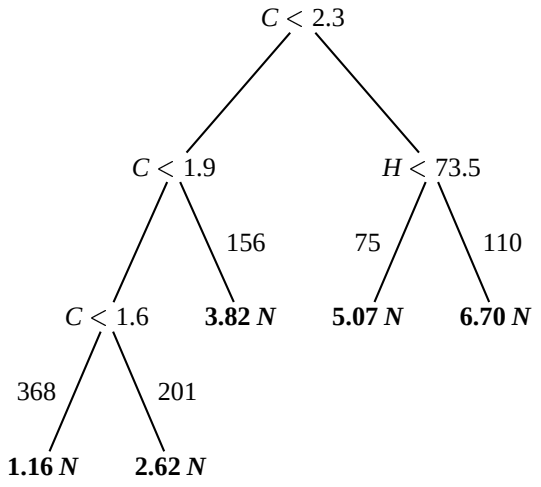

Fig. 7. Regression tree constructed for Northern Hemisphere Africa with Terra MODIS active fire data. Terminal nodes (leaves) are shown in boldface. The left fork is taken when the condition at a split is met. The numbers adjacent to each branch leading to a terminal node indicate the total number of observations assigned to the node during construction. To reduce clutter in the figure, subscripts have been dropped from the variables T_f (percent tree cover), H_f (percent herbaceous cover), B_f (percent bare ground), C_f (mean fire-pixel cluster size), and N_f (corrected fire pixel counts). Coefficients of N_f in terminal nodes have units of km^2/pixel.

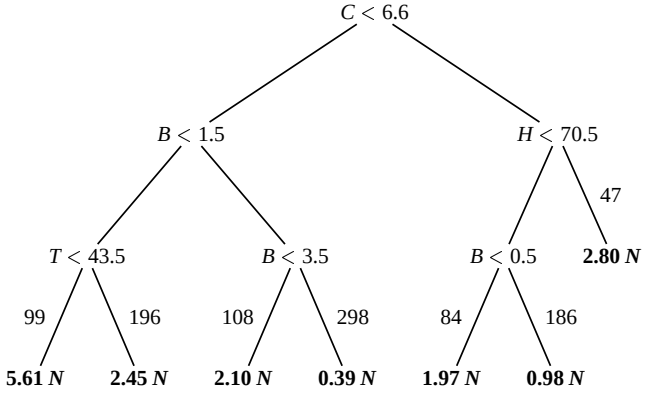

Fig. 8. Same as in Fig. 7 but for Boreal North America.

locations and time periods within the tropics for which a fixed correction factor, κ, will be applied to the burned area predicted via the regression trees. It may prove beneficial, in the future, to vary κ as a function of fire persistence in these areas, but we lack sufficient data to resolve this issue at present.

Our method of calculating fire persistence relied on the same subdivision of individual 1° grid cells into a fine grid of ∼1-km cells with which we identified fire-pixel clusters (Sect. 2.1). For each 1-km cell, we counted the number of days on which fires were detected during the particular

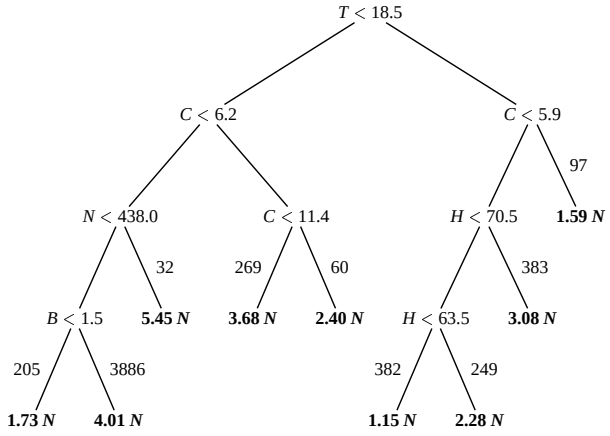

Fig. 9. Same as in Fig. 7 but for Australia.

month being processed. By averaging the number of days for all 1-km cells affected by fire, we computed the mean fire persistence (in days), denoted P_f, for each 1° grid cell on a monthly basis In Fig. 6 we show the climatological average of the monthly means (weighted by the number of fire pixels each month) from January 2001 through December 2004. Perhaps most obvious is the high persistence in Boreal North America, Boreal Asia, and the American Pacific Northwest, a consequence of the higher fuel loads and lower fire spread rates in the forested areas of these regions. Smaller patches of higher persistence are present in the Middle East region as well, where some residual gas-flare contamination remains in our MODIS fire data. Of most interest here, however, is the high persistence evident in areas of SH South America and Equatorial Asia in which deforestation is actively occurring. In these regions of slash-and-burn conversion, fire is a critical element of the deforestation process. It is also evident that high fire persistence does not occur in deforestation hot spots, such as Central Africa, where "slash-and-rot" conversion is commonplace (Achard et al., 1998).

Based on the previous discussion, it is relatively straightforward to identify grid cells within tropical rainforest for which application of the correction factor is appropriate; simple thresholds applied to tree cover and fire persistence, and restricted to the appropriate tropical regions, will suffice. Selection of an explicit value for κ, however, is more difficult. Although we lack the necessary data to adjust for cloud and canopy obscuration, we can at least use rough estimates to help correct for deforestation bias. Field observations suggest that forest clearings are typically one to ten times larger than the slash piles destined for burning, depending on whether the slashing was performed manually or mechanically (Douglas Morton and Wilfrid Schroeder, personal communication). We chose the geometric mean of these limits, yielding $\kappa \approx 3.2$. This is probably a more reasonable choice than the larger arithmetic mean as manual clearing is more prevalent in the tropics as a whole. In implementing the

correction, we identified those monthly grid cells within the tropics for which $T_f(i,t) \geq 50\%$ and $P_f(i,t) \geq 1.2$ days. For tropical grid cells satisfying these conditions, we multiplied the burned area predicted with the appropriate regression tree by κ.

4 Results

4.1 Regional regression trees

Regression trees were grown and pruned for each region; representative examples of the final result obtained for three of the fourteen regions are shown in Figs. 7 through 9. The number of terminal nodes in the final trees ranged from two (Europe and Equatorial Asia) to nine (Australia and SH South America). This wide range in size primarily reflects differences in the quantity of calibration data available for each region, and secondarily as an indication of the regional complexity in the burned-area/fire-count relationship.

Taking into consideration our previous discussion of the interpretation of regression trees grown from correlated variables (Sect. 3.2), we note that, in agreement with our preliminary analysis, both vegetation cover and fire-pixel cluster information play an important predictive role in the estimation of burned area. Considering the splits in all 14 regression trees together, 41 (68%) involved one of the three VCF variables, while 17 (28%) occurred on the mean fire-pixel cluster size. A detailed analysis of the final trees revealed that the reduction in deviance achieved by splitting on C_f was often much larger than that achieved by splitting on tree and herbaceous cover, indicating that the predictive utility of mean cluster size is not simply an artifact of its correlation with the VCF variables in the tropics. Only two splits (3%) occurred on N_f, primarily to deal with mild nonlinearities in the burned-area/fire-count relationship for SH Africa and Australia.

Figure 10 shows plots of predicted versus observed burned area for each region; the corresponding correlation coefficients are listed in Table 2. Regions showing the greatest agreement between predicted and observed burned area were Boreal Asia (r=0.94 with N=2104 observations), Central Asia (r=0.92, N=282), Europe (r=0.95, N=225), and Temperate North America (r=0.98, N=982); for these cases the predictions of area burned are comparatively accurate and precise. The region having poorest agreement was SH South America (r=0.56, N=4034), where predicted values of burned area suffer from large random and systematic errors. This is, at least in part, probably a consequence of the lower quality in 500-m burned area maps available for this region (Sect. 2.2). The remaining nine regions lie between these two extremes, yielding comparatively accurate, but imprecise, estimates of monthly burned area.

4.2 Uncertainties

Since we ultimately intend to use the regression trees to produce global monthly burned area estimates for input into coarse resolution models, it is important that uncertainty estimates be provided. A natural approach for quantifying these uncertainties would be through prediction confidence intervals computed for the fit of Eq. (2) within each terminal node of the regression tree. However, this approach is problematic in practice since the variance in burned area is not constant but instead increases as one considers larger burned areas. This behavior is characteristic of many physical variables and is referred to as *heteroskedasticity* in the statistical literature. Heteroskedasticity violates the constant-variance assumption of ordinary least squares fitting and, if ignored, may lead to inaccurate statistical error estimates for the fit.

Methods for dealing with heteroskedastic variables (weighted least squares, nonlinear data transformation) increase the influence of low-variance observations on the fit, while simultaneously decreasing the influence of high-variance observations. For our purposes this is undesirable. While it is true that the observations of very small burned areas have very low variability in an absolute sense (but very *high* variability in a relative sense), it is also true that these points are usually of less interest to most users. (An important exception is the burning of aggregated forest slash. As discussed in Sect. 3.3, such fires can consume prodigious quantities of biomass yet leave a very small burn scar.) In terms of emissions and land cover conversion, it is the observations of larger burns (which have high absolute variability but low relative variability) that are generally most important. Compensating for heteroskedasticity will therefore have the undesirable effect of assigning the greatest importance to the observations of least interest to us; this will in turn introduce biases (usually downward) in the predictions of large burned areas.

Given the above issues, we did not correct for heteroskedasticity when fitting Eq. (2) but adopted an alternative approach for estimating the uncertainties in our burned area estimates. For the fit of Eq. (2) within each terminal node of the regression tree, we regressed the square of the residuals (i.e. the variance) against fire pixel counts. The square root of the variance predicted by this supplementary fit then provided a one-standard-deviation ("one-sigma") uncertainty estimate for all future predictions emanating from the terminal node.

5 Multi-Year burned area estimates

We applied the regional regression trees to the entire archive of high-quality Terra MODIS data to produce a monthly global burned area data set spanning November 2000 through mid-2005. Using these data, we calculated the 2001–2004 mean monthly area burned and the associated uncertainties

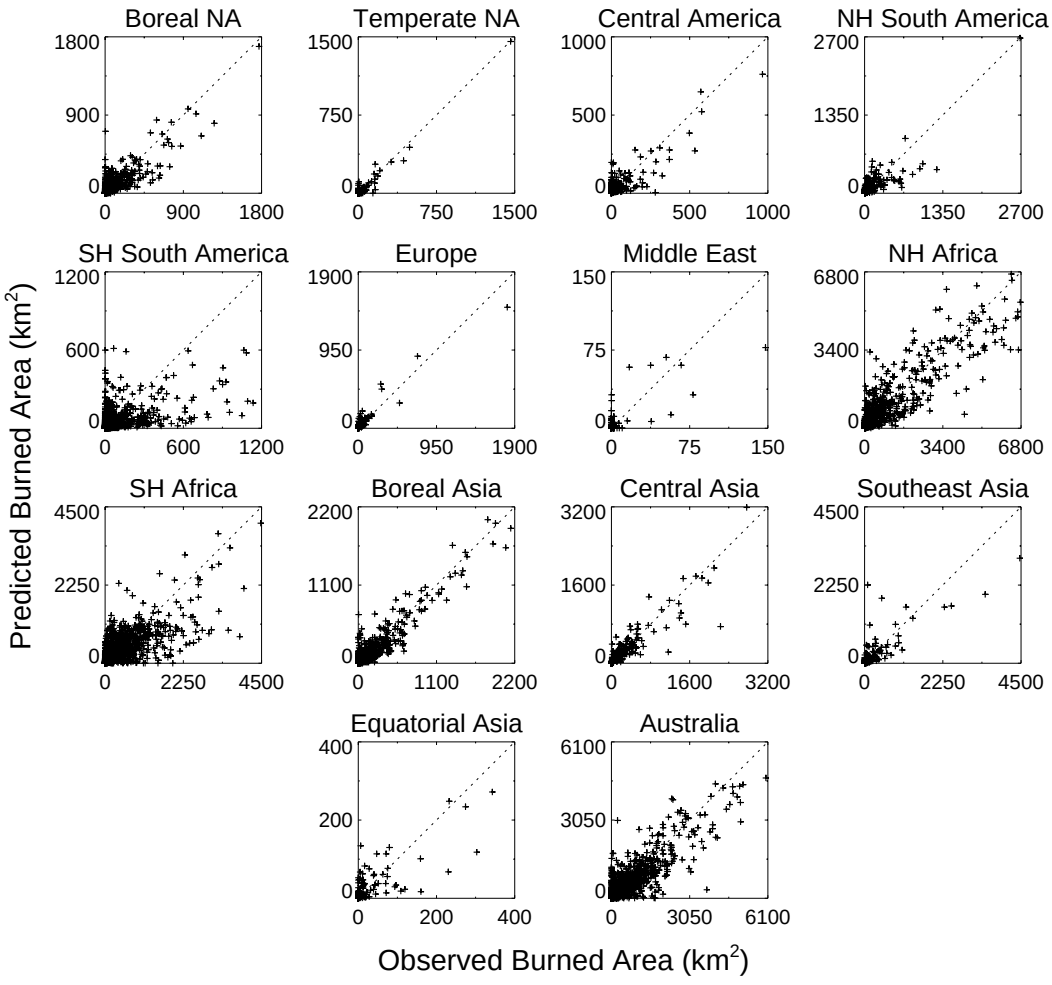

Fig. 10. Scatter plots of burned area predicted by regional regression trees vs. "true" burned area derived from 500-m burned area maps. Axes show area in km^2.

(Fig. 11). In propagating the monthly uncertainties we assumed they were random and independent, and hence added these in quadrature. (Given the large systematic errors noted earlier for SH South America, these estimates should be considered lower bounds in this region. The uncertainties in mean annual burned area suggested by Fig. 11 therefore probably underestimate the true error.) In general, the absolute uncertainties for regions characterized by large burned areas ($\geq\sim$1000 km^2/year) correspond to relative errors of 10% to 40%. In regions characterized by relatively small burned areas ($\leq$100 km^2/year), the absolute uncertainties typically correspond to much higher relative errors of 50% to 100%.

In Table 3 we show the annual area burned within each region for the years 2001–2004. The most extensive burning consistently occurred in northern hemisphere (NH) Africa, with well over 10^6 km^2 burned in this region each year. Over this four-year period substantial interannual variability – here arbitrarily defined as having at least one year of burned area varying by more than 50% of the four-year mean – occurred in Boreal North America, Boreal Asia, Equatorial Asia, and Australia. The interannual variability of burned area in NH Africa exceeded the annual area burned in all other regions, except SH Africa and Australia, over all four years. Taken together, the total area burned in northern- and southern-hemisphere Africa and Australia from 2001–2004 comprised 80% of the total area burned globally.

6 Evaluation

Rigorous validation of our global burned area data set requires independent, ground-truth quality maps of burn scars from representative locations over the entire globe. At present there is a paucity of such data, especially data that encompasses the monthly, 1° scale of our estimates. For example, global validation using direct estimates of burned area from Landsat imagery is limited by high data volumes and

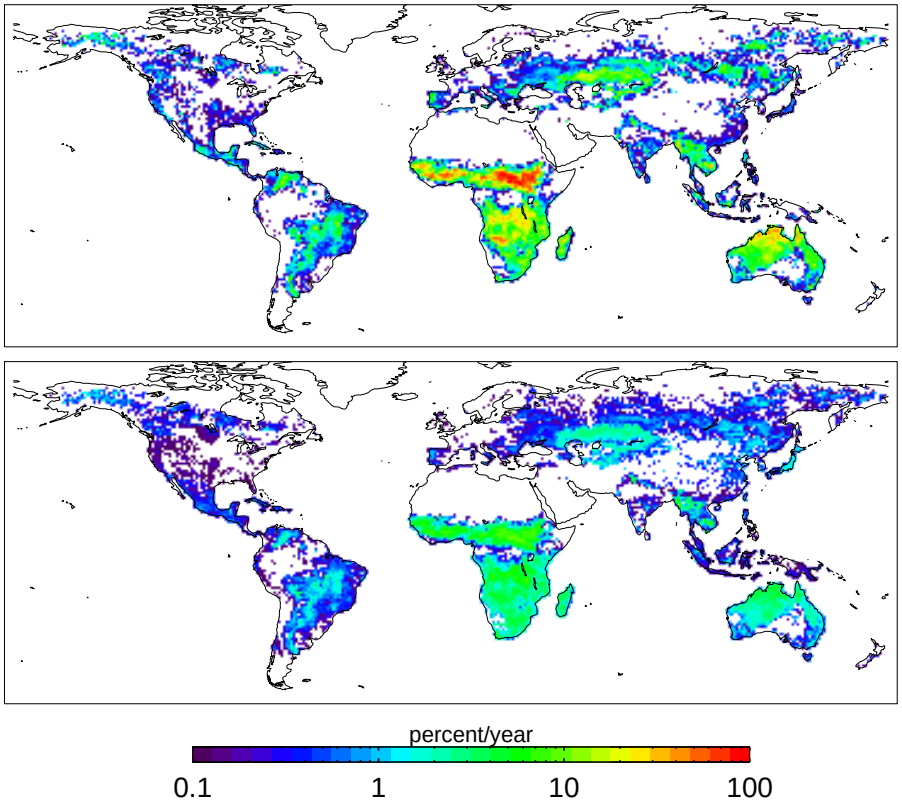

Fig. 11. 2001–2004 mean annual burned area derived from Terra MODIS active fire observations (top), and accompanying one-sigma uncertainties (bottom), expressed as the fraction of each grid cell that burns each year. One-sigma uncertainties were obtained by adding our spatially-explicit, monthly uncertainty estimates (assumed to be independent and random) in quadrature (Sect. 5). The estimation of monthly uncertainties is described in Sect. 4.2.

Table 3. 2001–2004 estimated annual area burned for the regions used in the study, with the mean of the relative errors (MRE) for the individual years shown in the rightmost column.

	Area Burned ($\times 10^4$ km^2=Mha)				
Region	2001	2002	2003	2004	MRE (%)
Boreal North America	0.4	2.6	2.3	4.0	8
Temperate North America	1.4	1.7	1.5	1.2	4
Central America	1.8	2.2	2.9	1.8	8
NH South America	4.4	3.6	4.8	3.8	6
SH South America	12.4	12.7	10.8	13.4	5
Europe	2.9	1.6	2.6	1.9	4
Middle East	0.6	0.5	0.4	0.4	8
NH Africa	153.2	135.2	125.5	129.8	2
SH Africa	84.0	82.4	79.6	75.3	3
Boreal Asia	6.3	9.3	14.5	4.9	3
Central Asia	16.5	26.7	17.1	18.9	4
Southeast Asia	10.8	10.2	8.4	16.1	7
Equatorial Asia	0.8	3.4	1.4	2.9	9
Australia	78.7	58.9	24.8	44.9	2

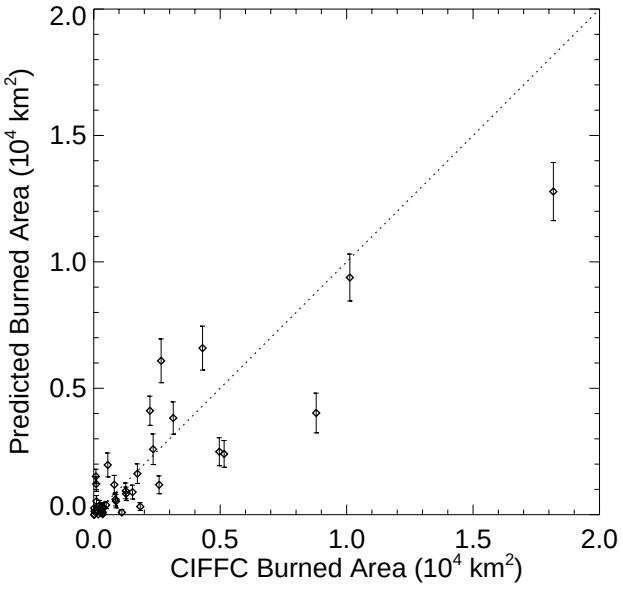

Fig. 12. Burned area predicted by regression tree each year for individual Canadian provinces during 2001–2004, versus annual provincial totals compiled by the Canadian Interagency Forest Fire Centre (http://www.ciffc.ca/). Error bars represent one-sigma uncertainties in predicted values.

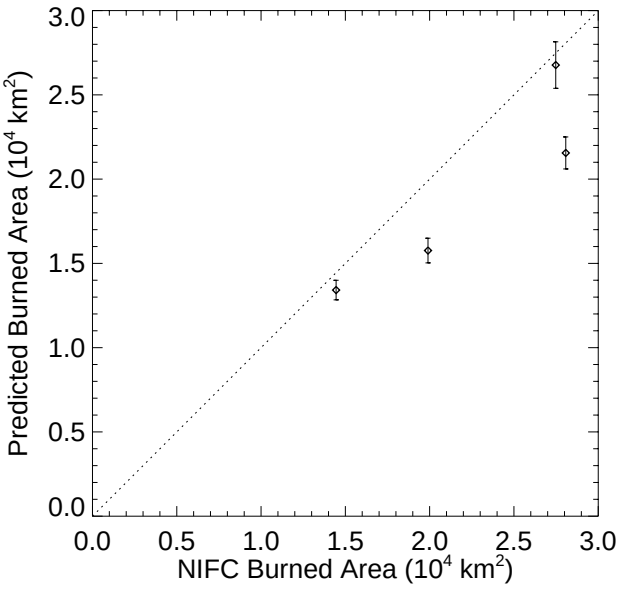

Fig. 13. Total predicted burned area within the United States for the years 2001–2004 versus annual totals compiled by the National Interagency Fire Center (http://www.nifc.gov/stats/index.html). Error bars represent one-sigma uncertainties in predicted values.

temporal discontinuities. A single Landsat scene provides coverage over an area approximately 180 km by 180 km in size, spanning in entirety at most a single 1° grid cell. When combined with the 16-day Landsat repeat cycle, it is difficult to unambiguously assign burned area to a specific calendar month. A practical (but more limited) alternative, which we describe here, is to compare our estimates of burned area to existing independent inventories. Results are summarized by region.

6.1 Canada

We compared our burned area estimates to independent estimates compiled by the Canadian Interagency Forest Fire Centre (CIFFC). These data are provided on a yearly basis from 2001–2004 for nine Canadian provinces (British Columbia, Alberta, Manitoba, Newfoundland and Labrador, Northwest Territories, Ontario, Quebec, Saskatchewan, and the Yukon Territories). A plot of predicted versus CIFFC burned area (Fig. 12) shows a very strong linear relationship (slope=0.70, r=0.89, $p \ll 0.001$), with some degree of underestimation for very large burned areas. This bias might in part be explained by the fact that ground-based and aerial surveys often record only the outermost perimeter of burn scars.

6.2 United States

We compared annual nationwide burned area statistics (including Alaska) compiled by the National Interagency Fire Center (NIFC) for the years 2001–2004 (Fig. 13). The two data sets are strongly linearly related (slope=0.83, r=0.91, p=0.093), although there is again a modest bias towards underestimation in large fire years. As with Canada, this bias might be partly explained by the manner in which large burns are surveyed. Considering the NIFC statistics as truth, the mean absolute percent error (MAPE) of our estimates is 13%.

6.3 Russia

We compared the total area burned in Russia for 2001 and 2002 to estimates produced by Sukhinin et al. (2004) from satellite data (Table 4). For 2001, a year of somewhat lower fire activity, our estimates were about 26% larger. For 2002 both totals agree to within a few tenths of a percent.

6.4 World

Our estimates of the total global annual area burned calculated for the years 2001–2004 range from a low of 2.97 million km^2 in 2003 to a high of 3.74 million km^2 in 2001. Although this does not qualify as an evaluation, we compared these results to the GBA2000 and GLOBSCAR products available for the year 2000 (Table 5). Despite the fact that we are comparing different years, our annual totals are only 0.3% to 27% higher than the total obtained with GBA2000.

Table 4. Comparison of predicted annual area burned in Russia with estimates of Sukhinin et al. (2004).

Year	Area Burned ($\times 10^4$ km^2=Mha) This study	Sukhinin et al. (2004)
2001	9.6	7.56
2002	12.1	12.1
2003	16.0	–
2004	6.9	–

Table 5. Comparison of global annual area burned obtained from the GBA2000 (Tansey et al., 2004) and GLOBSCAR (Simon et al., 2004) data sets, and estimated using the calibration approach described in this paper.

Source	Year	Area Burned ($\times 10^6$ km^2)
GBA2000	2000	2.93
GLOBSCAR	2000	1.94
This study	2001	3.74
	2002	3.51
	2003	2.97
	2004	3.19
	2001–2004 mean	3.35
GBA2000	Nov–Dec 2000	0.69
GLOBSCAR	Nov–Dec 2000	0.37
This study	Nov–Dec 2000	0.91

They are, however, substantially above the annual total obtained from GLOBSCAR, by a minimum of 51% and as much as 92%. Confining the comparison to November and December 2000, the only time period during which all three data sets overlap, our total burned area (0.91 million km^2) is about 32% higher than that of GBA2000, and 144% higher than that of GLOBSCAR. Thus, even if the 2002–2003 period corresponding to the weak El-Niño Southern Oscillation is not considered, the global burnt area estimate derived here is significantly higher than the GLOBSCAR estimate.

7 Application to other sensors

Clearly any active-fire calibration technique must be tuned to a specific sensor to compensate for differences in the characteristics of the sensor (e.g., spatial resolution), as well as the temporal sampling afforded by the platform on which the sensor resides. The latter issue is especially relevant because an increase in the rate at which "snapshots" of an active fire are made will generally increase the correspondence between a map of cumulative fire pixels and the spatial extent of the burn scar. It is instructive, therefore, to consider the effective temporal sampling rate of a particular instrument in gauging its suitability for providing burned area estimates via active-fire calibration. For a sensor on board a satellite having a polar or precessing orbit, a convenient measure of this sampling frequency is the daily equatorial coverage, denoted c_{eq}, which is simply the fraction of the Equator imaged by the sensor each day, irrespective of exactly where along the Equator the imaging occurs. For a swath width of w, this quantity is given by

$$c_{\mathrm{eq}} = \frac{w}{\pi R_E \sin\gamma}\left(\frac{24\,\mathrm{h}}{T}\right), \quad (4)$$

where R_E is the radius of the Earth, γ is the orbit inclination, and T is the orbital period (in hours). Values of the daily equatorial coverage for the MODIS, VIRS, and ATSR instruments (normalized to that of MODIS) are listed in Table 6. For the ATSR we explicitly considered the effect of restricting observations to nighttime overpasses, which halves the coverage predicted by Eq. (4), since the ATSR nighttime fire product is restricted in this manner.

Table 6. Swath width (w), orbit inclination angle (γ), and temporal sampling frequency (relative to that of the MODIS instruments) at the Equator for three sensors on-board operational satellites.

Sensor	w (km)	γ (°)	Relative Sampling Frequency
MODIS[1]	2330	98	1
VIRS[2]	830	35	0.66
ATSR[3]	512	99	0.22 (day)
			0.11 (night)

[1] Values applicable to Terra and Aqua MODIS instruments.
[2] Post August 2001 orbit boost.
[3] Entries are also applicable to the Advanced ATSR (AATSR).

Comparing the VIRS and MODIS instruments, the former provides about 34% less equatorial coverage; although the orbit inclination of the TRMM satellite on which VIRS resides increases the coverage by a factor of about two, the VIRS swath is narrower by a factor of about three. This suggests that a calibration-based technique based on VIRS active fire data is likely to yield lower quality estimates of burned area than for MODIS. More robust estimates are expected at higher subtropical latitudes where the VIRS sampling frequency increases by about a factor of three (Giglio et al., 2003b). However, the coarser VIRS spatial resolution (2.5 km) will probably degrade the quality of any calibration relationship involving cluster size, regardless of latitude, since VIRS fire-pixel clusters, being composed of larger areal units, provide a narrower range of unique values over which to discriminate different burned areas.

A comparison between MODIS and ATSR is more straightforward since these instruments have virtually

identical spatial resolution. Given the narrower swath, combined with the need to restrict fire observations to nighttime overpasses, the ATSR provides nearly an order of magnitude fewer opportunities to record active fire activity at the Equator. In the context of our calibration approach, this is equivalent to discarding nearly 90% of the MODIS active fire pixels recorded in each grid cell each month and repeating the calibration procedure. In addition, clusters of cumulative ATSR fire pixels are also likely to be smaller and more fragmented since the time interval between successive satellite overpasses is about ten times longer, thus reducing the utility of cluster-related information as a splitting variable. The combination of these factors is likely to yield larger uncertainties in estimates of burned area produced by calibrating ATSR active fire observations. Nevertheless, the ATSR has been shown to provide very useful estimates of the seasonal and interannual variability in burned area (e.g., Schultz, 2002).

8 Conclusions

We have presented a method for estimating monthly burned area globally at 1° spatial resolution using Terra MODIS active fire observations and ancillary vegetation cover information. Using regional regression trees, these data were calibrated to burned area estimates derived from 500-m MODIS imagery based on the conventional assumption that burned area is proportional to counts of fire pixels under specific conditions. Traditionally, the constant of proportionality (α) has either been held fixed, or adjusted based on a single vegetation-related parameter. Neither practice is satisfactory at a global scale. We propose a more flexible approach in which α is permitted to vary as a function of both tree and herbaceous vegetation cover (or alternatively bare ground fraction), and the mean size of monthly cumulative fire-pixel clusters within each 1° grid cell. Though we found this to be usually unnecessary, we also allowed α to vary with fire pixel counts to accommodate slight deviations from the assumption of linearity. The exact form of the functional dependence of α on these predictive variables was not specified a priori, but was constructed through recursive partitioning and expressed in terms of the splits and leaves of a regression tree. In addition to their considerable flexibility, regression trees offer the advantage of readily accommodating additional explanatory variables on a trial basis.

Recognizing limits in our ability to measure burned area in closed canopy tropical forests, we used information about monthly fire persistence and tree cover to identify locations and time periods within the tropics requiring the application of a fixed correction factor to the burned area predictions obtained from the regression trees.

Regions showing good agreement between predicted and observed burned area included Boreal Asia, Central Asia, Europe, and Temperate North America, where the estimates produced by the regression trees were relatively accurate and precise. Poorest agreement was found for SH South America, where predicted values of burned area are both inaccurate and imprecise. The poor result obtained in this region is most likely a consequence of multiple factors that include extremely persistent cloud cover and a degradation in the quality of the 500-m burned area maps used for calibration. Agreement in the nine remaining regions fall between these two extremes, yielding comparatively accurate, but less precise, estimates of monthly burned area.

We used the regional regression trees to produce multi-year, global burned area estimates on a monthly basis from the current archive of Terra MODIS active fire data. Annual totals derived from these data showed good agreement with independent annual estimates available for nine Canadian provinces, the continental United States, and Russia. Using these data, we estimated the global annual burned area for the years 2001–2004 to vary between 2.97×10^6 and 3.74×10^6 km^2, with the maximum occurring in 2001. The most extensive burning consistently occurred in NH Africa, with well over 10^6 km^2 burned in this region each year. Over this four-year period significant interannual variability occured in Boreal North America, Boreal Asia, Equatorial Asia, and Australia. Taken together, the total area burned in northern- and southern-hemisphere Africa and Australia from 2001–2004 comprised 80% of the total area burned globally.

We reiterate that we are not promoting our regression-tree approach (or, indeed, any active-fire calibration approach) as a substitute for burned area maps generated from direct observations of burn scars. Rather, for some applications, statistical coarse-resolution burned area estimates derived from MODIS active fire observations can serve as a useful interim product until long-term burned area data sets become available.

Appendix A Burned area detection algorithm

Burned area maps were produced using a prototype algorithm (Giglio et al., 2006[1]) that uses the 500-m MODIS atmospherically-corrected Level 2G surface reflectance product (Vermote et al., 2002), the MODIS Level 3 daily active fire products (Justice et al., 2002), and the MODIS Level 3 96-day Land Cover Product (Friedl et al., 2002). The algorithm, which is a major extension of an earlier method proposed by Roy et al. (1999), detects persistent changes in a daily vegetation-index (VI) time series derived from MODIS band 5 (1.2 μm) and band 7 (2.1 μm) surface reflectances, respectively denoted ρ_5 and ρ_7, where

$$\mathrm{VI} = \frac{\rho_5 - \rho_7}{\rho_5 + \rho_7}. \quad \text{(A1)}$$

[1]Giglio, L., Descloitres, J., Loboda, T., Csiszar, I., and Kendall, J.: An Active-Fire Based Burned Area Mapping Algorithm for the MODIS Sensor, in preparation, 2006.

This index shows a significant decrease following a burn, and provides somewhat better discrimination of burned areas than the more commonly used Normalized Burn Ratio (NBR), an index defined similarly but with Landsat Thematic Mapper (TM) bands 4 (0.83 μm) and 7 (2.2 μm) (Miller and Yool, 2002).

The general detection approach is to first derive a summary map of persistent change from the VI time series, and then use spatial and temporal active fire information to guide the statistical characterization of burn-related and non burn-related change within the scene. This information is used to estimate probabilistic thresholds suitable for classifying the scene into burned and unburned pixels. The approach ultimately identifies the date of burn, to the nearest day, for pixels within individual MODIS Level 3 tiles (Wolfe et al., 1998) at 500-m spatial resolution.

A1. Composite change summary

The algorithm first examines the daily VI time series by considering observations within two adjacent sliding temporal windows of duration W=10 days; these windows are referred to as the *candidate pre-burn* and *candidate post-burn* windows, respectively. Within the k'th candidate pre-burn window, the trimmed mean [$\mathrm{VI}_{\mathrm{pre}}(k)$] and trimmed standard deviation [$\sigma_{\mathrm{pre}}(k)$] of all observations are computed. Statistics for the k'th candidate post-burn window are similarly computed and denoted $\mathrm{VI}_{\mathrm{post}}(k)$ and $\sigma_{\mathrm{post}}(k)$. The index k references the position within the daily time series on which the sliding windows are aligned (incrementing k moves both windows forward in time by one day). The time series of active fire observations for the pixel under consideration is also examined, and the occurrence of any such pixels in the time series is flagged.

A measure of temporal separability $S(k)$, defined as

$$S(k) = \frac{\Delta\mathrm{VI}(k)}{\sigma_{\mathrm{pre}}(k) + \sigma_{\mathrm{post}}(k)}, \quad \text{(A2)}$$

where $\Delta\mathrm{VI}(k)=\mathrm{VI}_{\mathrm{pre}}(k)-\mathrm{VI}_{\mathrm{post}}(k)$, is evaluated for all k. For each pixel the maximum separability $\max(S(k))=S(k_{\max})$ is identified. The date associated with the maximum change is the midpoint of the interval between the last observation in the pre-burn window and the first observation in the post-burn window. The time series for each pixel of the MODIS tile is processed, yielding composite images of $\Delta\mathrm{VI}(k_{\max})$ and $\mathrm{VI}_{\mathrm{post}}(k_{\max})$, and a composite map of active fire pixels detected during the time period being processed.

A2. Identification of training samples

Having generated the composite imagery just described, a simple procedure is used to extract representative burned and unburned samples within each land cover class l present within the MODIS tile. To identify probable burned pixels, the composite active fire map is morphologically eroded. This eliminates the smallest fire-pixel clusters, which are less likely to be accompanied by a detectable burn scar (cf. Fig. 1). Burned training samples of $\Delta\mathrm{VI}(k_{\max})$ and $\mathrm{VI}_{\mathrm{post}}(k_{\max})$ are drawn from those pixels remaining in the eroded fire mask, and are partitioned by land cover class. In a complementary manner, dilating the composite active fire map provides a mask of pixels that are unlikely to have burned during the time period being processed, and training samples of unburned pixels for each land cover class are thus identified. Here more care is required since the active fire map will often greatly under-represent the spatial extent of large burns, so the radius of the dilation kernel is increased in proportion to the size of individual fire-pixel clusters. This is again consistent with the empirical observation that large burns tend to be accompanied by large clusters of active fire pixels, and vice versa.

A3. Classification of unambiguous burned and unburned pixels

The training samples extracted in the previous step are now used to derive conservative dynamic thresholds to classify unambiguous burned and unburned pixels. Pixels with land cover class l for which $\Delta\mathrm{VI}(k_{\max})$ is less than the upper quartile of $\Delta\mathrm{VI}(k_{\max})$ for all unburned training pixels in land cover class l are immediately labeled as *unburned*. Pixels with land cover class l for which $\Delta\mathrm{VI}(k_{\max})$ is greater than the upper quartile of $\Delta\mathrm{VI}(k_{\max})$, and $\mathrm{VI}_{\mathrm{post}}(k_{\max})$ is less than the lower quartile of $\mathrm{VI}_{\mathrm{post}}(k_{\max})$ for all burned training pixels in land cover class l are immediately classified as *burned*. Prior to performing this step an initial separability test is performed for each land cover class: if the distributions of $\Delta\mathrm{VI}(k_{\max})$ for the burned and unburned training samples show excessive overlap, all pixels within land cover class l are immediately classified as *unburned* and the quartile tests are not performed.

This step typically provides a final classification for 60% to 80% of all image pixels, resulting in a large reduction in the computational effort needed to process pixels in the remainder of the scene.

A4. Classification of remaining pixels

Following the inital labeling of obvious burned and unburned pixels, one of two different approaches are then used to label the remaining unclassified pixels depending on the geographic region and time period being processed.

A4.1. Region growing

In high-latitude regions having shorter MODIS revisit periods, and at lower latitudes where the majority of the area burned is dominated by large burn scars ($>\sim$100 km^2), region-growing is used to identify the remaining burned pixels within the MODIS tile being processed. Clusters of active fire pixels derived from the composite active-fire mask

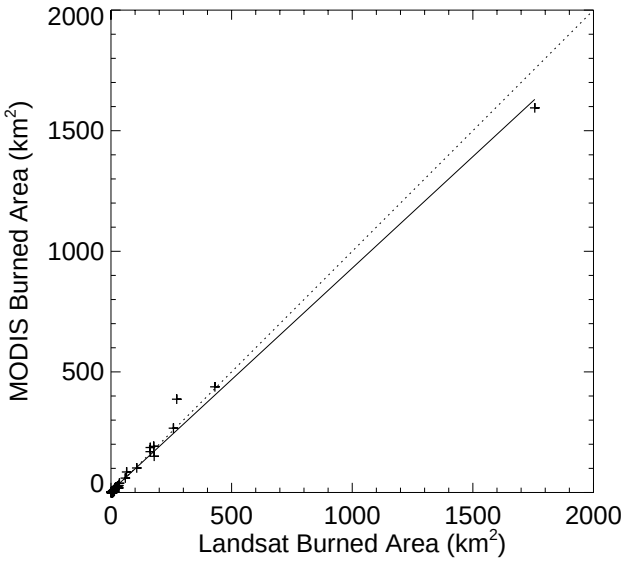

Fig. A1. Areas of 39 individual burn scars in Russia during 2001 and 2002 as mapped by 500-m burned area detection algorithm versus ground truth derived manually from Landsat imagery. The slope and intercept of the solid black regression line are 0.925 and 5.5 km^2, respectively, with a correlation of 0.996.

are used as seeds to iteratively "fill in" the surrounding burn scar. The values of $\Delta\mathrm{VI}(k_{\max})$ and $\mathrm{VI}_{\mathrm{post}}(k_{\max})$ for unclassified candidate pixels adjacent to a seed pixel are compared to the statistical distributions derived from the burned and unburned training samples. Bayes' rule is applied to estimate a posteriori probabilities for the candidate pixel and select the class to which it will be assigned. Additional temporal constraints are applied based on the burn dates of the parent seeds. Pixels classified as *burned* during the current iteration become seeds in the subsequent iteration. The process continues until no new seeds are found.

A4.2. Contextual classification

In regions where the temporal sampling of active fires is less frequent, or relatively small burns ($<\sim$10 km^2) are abundant, an alternate classification approach is required. Region growing under these conditions will usually result in large omission errors since many smaller burns will lack active-fire seed pixels, and a contextual classifier is employed instead. In addition to using the spectral information available for the training pixels and the entire scene, the contextual classifier exploits the fact that both burned and unburned pixels tend to occur near pixels having the same burned or unburned status. Here, a Markov Random Field (MRF), a standard approach for modeling such spatial behavior, is employed. The MRF is defined by a probability density function that encodes the likelihood that a pixel is burned based on the state of its eight immediate neighbors. This information is incorporated with the spectral information provided by $\Delta\mathrm{VI}(k_{\max})$ and $\mathrm{VI}_{\mathrm{post}}(k_{\max})$, and each pixel not classified during step A.3 is assigned the class (*burned* or *unburned*) having the maximum a posteriori (MAP) probability. Estimation of these probabilities is performed using the iterated conditional modes (ICM) method described by Besag (1986). ICM requires a rough preliminary classification to initiate the iterative process, and for this a naive Bayes classifier is used. The preliminary class assigned to each pixel within land cover class l is that which has the highest a posteriori probability estimated using Bayes' rule:

$$p_l(\mathrm{burned}|\Delta\mathrm{VI},\,\mathrm{VI}_{\mathrm{post}}) = \frac{p_l(\Delta\mathrm{VI}|\mathrm{burned})\,p_l(\mathrm{VI}_{\mathrm{post}}|\mathrm{burned})\,p_l(\mathrm{burned})}{p_l(\mathrm{VI},\,\Delta\mathrm{VI}_{\mathrm{post}})} \quad \text{(A3)}$$

$$p_l(\mathrm{unburned}|\Delta\mathrm{VI},\,\mathrm{VI}_{\mathrm{post}}) = \frac{p_l(\Delta\mathrm{VI}|\mathrm{unburned})\,p_l(\mathrm{VI}_{\mathrm{post}}|\mathrm{unburned})\,p_l(\mathrm{unburned})}{p_l(\mathrm{VI},\,\Delta\mathrm{VI}_{\mathrm{post}})} \quad \text{(A4)}$$

(Note that the index $k_{\max}$ has been dropped in the above equations to reduce clutter.) Again making use of information provided by active fire observations, the a priori probability of a burned pixel $p_l(\mathrm{burned})$ in land cover class l is estimated as the fraction of all pixels of land cover class l within the scene in which an active fire was detected during the time period being processed. From this $p_l(\mathrm{unburned})=1-p_l(\mathrm{burned})$.

A5. Validation

At present, validation of the 500-m burned area maps has been limited to Russia through comparison with maps generated manually from high resolution satellite imagery (Loboda and Csiszar, 2004). Using 20 Landsat scenes acquired between 17 August 2001 and 19 August 2002, Loboda and Csiszar found good agreement between the burn scars mapped by the algorithm with those traced manually (Fig. A1). Validation efforts are currently being extended to Australia and South America.

Acknowledgements. We thank D. Morton and W. Schroeder for helpful technical discussions, and two anonymous reviewers for their helpful comments. This work was supported by NASA grants NNG04GD89G and NNG04GK49G, and NASA contract NNH04CC22C. The MODIS burned area mapping algorithm summarized in the Appendix was developed by L. Giglio, J. Descloitres, and J. Kendall (SSAI). The resulting burned area maps are being validated by T. Loboda, I. Csiszar, D. Morton, S. Trigg (UMD), and K. Murphy (SSAI).

Edited by: F. J. Dentener

References

Achard, F., Eva, H. D., Glinni, A., Mayaux, P., Stibig, H.-J., and Richards, T.: Identification of Deforestation Hot Spot Areas in the Humid Tropics, TREES Publications Series B, Research Report No. 4, EUR 18079 EN, European Commission, Luxembourg, 1998.

Arino, O. and Rosaz, J.: 1997 and 1998 world ATSR fire atlas using ERS-2 ATSR-2 data, Proceedings of the Joint Fire Science Conference, Boise, Idaho, 15–17 June 1999, edited by: Neuenschwander, L. F., Ryan, K. C., and Golberg, G. E., Boise: University of Idaho and the International Association of Wildland Fire, 177–182, 1999.

Besag, J. E.: On the statistical analysis of dirty pictures, J. Royal Stat. Soc., Ser. B, 48, 259–302, 1986.

Boschetti, L., Eva, H. D., Brivio, P. A., and Grégoire, J.-M.: Lessons to be learned from the comparison of three satellite-derived biomass burning products, Geophys. Res. Lett., 31, L21501, doi:10.1029/2004GL021229, 2004.

Breiman, L., Friedman, J. A., Olshen, R. A., and Stone, C. J.: Classification and Regression Trees, Boca Raton, Chapman & Hall/CRC, 1984.

Breiman, L. and Meisel, W. S.: General estimates of the intrinsic variability of data in nonlinear regression models, J. Amer. Stat. Assoc., 71, 301–307, 1976.

Duncan, B. N., Martin, R. V., Staudt, A. C., Yevich, R., and Logan, J. A.: Interannual and seasonal variability of biomass burning emissions constrained by satellite observations, J. Geophys. Res.-Atmos., 108(D2), 4100, doi:10.1029/2002JD002378, 2003.

Eva, H. and Lambin, E. F.: Remote sensing of biomass burning in tropical regions: sampling issues and multisensor approach, Rem. Sens. Environ., 64, 292–315, 1998.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.: Global land cover mapping from MODIS: algorithms and early results, Rem. Sens. Environ., 83, 287–302, 2002.

Generoso, S., Bréon, F.-M., Balkanski, Y., Boucher, O., and Schulz, M.: Improving the seasonal cycle and interannual variations of biomass burning aerosol sources, Atmos. Chem. Phys., 3, 1211–1222, 2003.

Giglio, L., Kendall, J. D., and Mack, R.: A multi-year active fire data set for the tropics derived from the TRMM VIRS, Int. J. Rem. Sens., 24, 4505–4525, 2003a.

Giglio, L., Pinzon, J., and Kasibhatla, P.: Comment on "Seasonal, intraseasonal, and interannual variability of global land fires and their effects on atmospheric aerosol distribution" by Y. Ji and E. Stocker, J. Geophys. Res.-Atmos., 108(D24), 4754, doi:10.1029/2003JD003548, 2003b.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C., and Sohlberg, R. A.: Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm, Earth Interactions, 7(10), doi:10.1175/1087-3562, 2003.

Hao, W.-M., Liu, M.-H., and Crutzen, P. J.: Estimates of annual and regional releases of CO_2 and other trace gases to the atmosphere from fires in the tropics, based on the FAO statistics for the period 1975-1980, in: Fire in the Tropical Biota: Ecosystem Processes and Global Challenges, edited by: Goldammer, J. G., pp. 400–462, Springer-Verlag, New York, 1990.

Heald, C. L., Jacob, D. J., Palmer, P. I., Evans, M. J., Sachse, G. W., Singh, H. B., and Blake, D. R.: Biomass burning emission inventory with daily resolution: Application to aircraft observations of Asian outflow, J. Geophys. Res.-Atmos., 108(D21), 8811, doi:10.1029/2002JD003082, 2003.

Justice, C. O., Giglio, L., Korontzi, S., Owens, J., Morisette, J., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F., and Kaufman, Y.: The MODIS fire products, Rem. Sens. Environ., 83, 244–262, 2002.

Kasischke, E. S., Hewson, J. H., Stocks, B., van der Werf, G., and Randerson, J.: The use of ATSR active fire counts for estimating relative patterns of biomass burning – a study from the boreal forest region, Geophys. Res. Lett., 30(18), 1969, doi:10.1029/2003GL017859, 2003.

Langmann, B. and Heil, A.: Release and dispersion of vegetation and peat fire emissions in the atmosphere over Indonesia 1997/1998, Atmos. Chem. Phys., 4, 2145–2160, 2004.

Loboda, T. and Csiszar, I.: Estimating burned area from AVHRR and MODIS: validation results and sources of error. Proceedings of the Current Aspects of Rem. Sens. Earth Space Conf., Moscow, Russia, 16–18 November, 2004.

Miller, J. D. and Yool, S. R.: Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data, Rem. Sens. Environ., 82, 481–496, 2002.

Pereira, J. M., Pereira, B. S., Barbosa, P., Stroppiana, D., Vasconcelos, M. J. P., and Grégoire, J.-M.: Satellite monitoring of fire in the EXPRESSO study area during the 1996 dry season experiment: Active fires, burnt area, and atmospheric emissions, J. Geophys. Res.-Atmos., 104(D23), 30 701–30 712, 1999.

Potter, C., Genovese, V. B., Klooster, S., Bobo, M., and Torregrosa, A.: Biomass burning losses of carbon estimated from ecosystem modeling and satellite data analysis for the Brazilian Amazon region, Atmos. Environ., 35, 1773–1781, 2001.

Prins, E. M., Feltz, J. M., Menzel, W. P., and Ward, D. E.: An overview of GOES-8 diurnal fire and smoke results for SCAR-B and the 1995 fire season in South America, J. Geophys. Res.-Atmos., 103(D24), 31 821–31 836, 1998.

Randriambelo, T., Baldy, S., Bessafi, M., Petit, M., and Despinoy, M.: An improved detection and characterization of active fires and smoke plumes in south-eastern Africa and Madagascar, Int. J. Rem. Sens., 19(14), 2623–2638, 1998.

Roy, D. P., Giglio, L., Kendall, J. D., and Justice, C. O.: A multitemporal active-fire based burn scar detection algorithm, Int. J. Rem. Sens., 20, 1031–1038, 1999.

Roy, D. P., Lewis, P. E., and Justice, C. O.: Burned area mapping using multi-temporal moderate spatial resolution data – a bi-directional reflectance model-based expectation approach, Rem. Sens. Environ., 83, 263–286, 2002.

Scholes, R. J., Kendall, J. D., and Justice, C. O.: The quantity of biomass burned in southern Africa, J. Geophys. Res.-Atmos., 101, 23 667–23 676, 1996.

Schultz, M.: On the use of ATSR fire count data to estimate the seasonal and interannual variability of vegetation fire emissions, Atmos. Chem. Phys., 2, 387–395, 2002.

Seiler, W. and Crutzen, P. J.: Estimates of gross and net fluxes of carbon between the biosphere and atmosphere from biomass burning, Climate Change, 2, 207–247, 1980.

Setzer, A. W. and Pereira, M. C.: Amazonia biomass burnings in 1987 and an estimate of their tropospheric emissions, Ambio,

20, 19–22, 1991.

Simon, M., Plummer, S., Fierens, F., Hoelzemann, J. J., and Arino, O.: Burnt area detection at global scale using ATSR-2: The GLOBSCAR products and their qualification, J. Geophys. Res.-Atmos., 109, D14S02, doi:10.1029/2003JD003622, 2004.

Streets, D. G., Yarber, K. F., Woo, J.-H., and Carmichael, G. R.: Biomass burning in Asia: Annual and seasonal estimates and atmospheric emissions, Global Biogeochem. Cycles, 17(4), 1099, doi:10.1029/2003GB002040, 2003.

Stroppiana, D., Brivio, P. A., and Grégoire, J.-M.: Modelling the impact of vegetation fires, detected from NOAA-AVHRR data, on tropospheric chemistry in tropical Africa, in: Biomass Burning and its Inter-Relationships with the Climate System, edited by: Innes, J. L., Beniston, M., and Verstraete, M. M., Dordrecht: Kluwer Academic Publishers, pp. 193–213, 2000.

Sukhinin, A. I., French, N. H. F., Kasischke, E. S., Hewson, J. H., Soja, A. J., Csiszar, I. A., Hyer, E. J., Loboda, T., Conard, S. G., Romasko, V. I., Pavlichenko, E. A., Miskiv, S. I., and Slinkina, O. A.: Satellite-based mapping of fires in Russia: New products for fire management and carbon cycle studies, Rem. Sens. Environ., 93, 546–564, 2004.

Tansey, K., Grégoire, J.-M., Stroppiana, D., Sousa, A., Silva, J., Pereira, J. M. C., Boschetti, L., Maggi, M., Brivio, P. A., Fraser, R., Flasse, S., Ershov, D., Binaghi, E., Graetz, D., and Peduzzi, P.: Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data, J. Geophys. Res.-Atmos., 109, D14S03, doi:10.1029/2003JD003598, 2004.

van der Werf, G. R., Randerson, J. T., Collatz, G. J., and Giglio, L.: Carbon emissions from fires in tropical and subtropical ecosystems, Global Change Biology, 9, 547–562, 2003.

van der Werf, G. R., Randerson, J. T., Collatz, G. J., Giglio, L., Kasibhatla, P. S., Arellano Jr., A. F., Olsen, S. C., and Kasischke, E. S.: Continental-scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period, Science, 303, 73–76, 2004.

Vermote, E. F., El Saleous, N. Z., and Justice, C. O.: Operational atmospheric correction of the MODIS data in the visible to middle infrared: First results, Rem. Sens. Environ., 83, 97–111, 2002.

Wolfe, R. E., Roy, D. P., and Vermote, E.: MODIS land data storage, gridding, and compositing methodology: Level 2 grid, IEEE Trans. Geosci. Rem. Sens., 36, 1324–1338, 1998.