

# What limits fire? An examination of drivers of burnt area in Southern Africa

SALLY ARCHIBALD\*, DAVID P. ROY†, BRIAN W. VAN WILGEN‡ and ROBERT J. SCHOLES\*

\*Natural Resources and the Environment, CSIR, PO Box 395, Pretoria 0001, South Africa, †Geographic Information Science Centre of Excellence, South Dakota State University, Brookings, SD 57007, USA, ‡Centre for Invasion Biology, Natural Resources and the Environment, CSIR, PO Box 320, Stellenbosch, South Africa,

## Abstract

The factors controlling the extent of fire in Africa south of the equator were investigated using moderate resolution (500 m) satellite-derived burned area maps and spatial data on the environmental factors thought to affect burnt area. A random forest regression tree procedure was used to determine the relative importance of each factor in explaining the burned area fraction and to address hypotheses concerned with human and climatic influences on the drivers of burnt area. The model explained 68% of the variance in burnt area. Tree cover, rainfall in the previous 2 years, and rainfall seasonality were the most important predictors. Human activities – represented by grazing, roads per unit area, population density, and cultivation fraction – were also shown to affect burnt area, but only in parts of the continent with specific climatic conditions, and often in ways counter to the prevailing wisdom that more human activity leads to more fire. The analysis found no indication that ignitions were limiting total burnt area on the continent, and most of the spatial variation was due to variation in fuel load and moisture. Split conditions from the regression tree identified (i) low rainfall regions, where fire is rare; (ii) regions where fire is under human control; and (iii) higher rainfall regions where burnt area is determined by rainfall seasonality. This study provides insights into the physical, climatic, and human drivers of fire and their relative importance across southern Africa, and represents the beginnings of a predictive framework for burnt area.

*Keywords:* climatic control, dry season, fire regimes, fuel, human control, ignitions, predictive model, regression tree, weather

*Received 10 October 2007; revised version received 6 February 2008 and accepted 10 June 2008*

## Introduction

A prolonged annual dry season combined with relatively rapid rates of fuel accumulation create conditions conducive to frequent vegetation fires in southern Africa. Indeed, fire is considered a major determinant of the ecology and distribution of Africa's savanna and grassland vegetation types (VanWilgen & Scholes, 1997; Higgins *et al.*, 2000; Bond *et al.*, 2005). The fraction of the landscape that burns varies greatly across the region. Understanding the causes of this variation is not simple because the propensity to burn is influenced by weather conditions, the presence of ignition sources, and the amount, type, and arrangement of the available fuel, all of which change across space and through time.

Satellite data products and spatially explicit environmental data are increasingly becoming available, and continental-scale information on many environmental factors, including fire, is now relatively easy to obtain. Quantitative methods can now be applied to address scientific questions that previously relied on anecdotal information or localised studies; in particular, how the various drivers of fire interact to result in characteristic fire regimes. These questions have become increasingly important in the context of global change. Fire is both an important determinant of vegetation community structure (Bond & Van Wilgen, 1996), and a globally significant source of greenhouse gas emissions (Patra *et al.*, 2005; Williams *et al.*, 2007). Fire patterns will almost certainly change in response to population, land use, and climatic changes in Africa (Boko *et al.*, 2007), and an understanding of the drivers of such changes will be needed to predict their consequences.

Correspondence: Sally Archibald, tel. +27 12 841 3487, fax +27 12 841 4322, e-mail: sarchibald@csir.co.za

A century of experimental and descriptive research has provided a good understanding of the physical factors controlling the ignition and spread of both grass-fuelled and forest fires (see in particular Byram, 1959; Luke & McArthur, 1978; Trollope *et al.*, 2002; Viegas, 2004). The challenge is to determine which factors dominate under different conditions, and to predict the outcome of the interaction of several, often antagonistic, influences. For example, areas of high population density have been shown to be associated with an increase in the number of fires (Keeley *et al.*, 1999), but increased population densities also result in more intensive land use, reduced fuel loads, and fragmentation of landscapes, which act to reduce the spread of fire (Frost, 1999). Similarly, although fuel production generally increases with rainfall, so too do factors with negative effects on fire, such as fuel moisture (Scholes *et al.*, 1996; Spessa *et al.*, 2005). Thus, it is the relative importance of contrasting influences acting under different circumstances that determines the characteristic fire regime of a particular region.

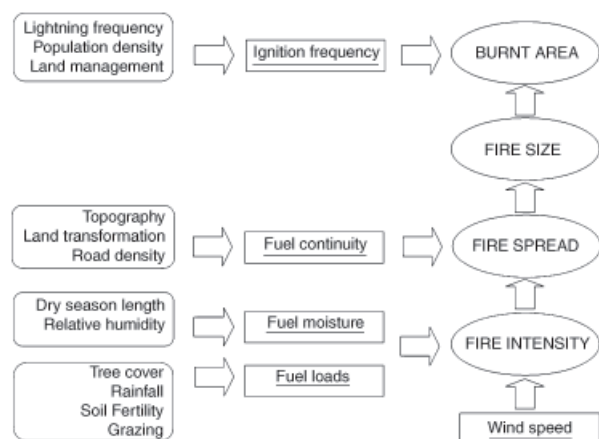
This paper aims to investigate the factors controlling the annual fraction of the landscape that burns across Africa south of the equator. The fire regime of an area is defined by several variables, including the intensity, season, and type of fire (Gill, 1975). The average period between fires, often confusingly referred to as the 'fire frequency' is fundamentally important. In this paper, we analyse the annual burnt area fraction, which is usually regarded as the reciprocal of the fire return time (Frost, 1999). Burned area fraction is also important in its own right. For example, it is necessary for calculating greenhouse gas emissions (Crutzen & Andreae, 1990).

Information on the relative importance of various environmental factors affecting burnt area in southern Africa is currently lacking. Further, there is little agreement over how much influence people have, relative to abiotic factors such as climate, in controlling fires (see Heyerdahl *et al.* (2001); Keeley & Fotheringham (2001); Moritz (2003); Dickson *et al.* (2006); Westerling *et al.* (2006) for a recent north American discussion on this). Human population data have been incorporated into mechanistic fire models (Thonicke *et al.*, submitted; Venevsky *et al.*, 2002), but in Africa fundamental assumptions on the extent to which human activities constrain or promote fire are yet to be tested over large areas. Rather than modelling fire processes mechanistically, we adopt a statistical modelling approach. This research provides new, continental scale insights into the physical, climatic, and human drivers of fire and the relative importance of these across southern Africa.

## Hypotheses

The majority of fires in southern Africa are surface fires, fuelled by grass and litter. Given sufficient rainfall, this fuel can regrow rapidly after the fire, and can cure and be ready to burn after only a few weeks of dry weather (Stott, 2000). Current theoretical understanding of the drivers of fire in southern Africa is synthesised in Fig. 1. Rainfall and soil nutrients positively affect grass production, and therefore grass fuels; however, high tree covers and high grazing pressure will reduce grass fuels (Trollope, 1984; VanWilgen & Scholes, 1997). The duration of the dry season will determine the amount of time that fuel is dry and available to burn and, combined with weather conditions on the day of burning, determine fuel moisture (Spessa *et al.*, 2005; Russell-Smith *et al.*, 2007). Variations in lightning frequencies and human population densities are likely to affect ignition frequencies (Keeley *et al.*, 1999), and land management may affect both the ignition and suppression of fires (Frost, 1999). Fuel continuity is impacted both by the landscape morphology – highly dissected, variable landscapes will prevent the spread of large fires (Dickson *et al.*, 2006; Russell-Smith *et al.*, 2007); and by human activities – building of roads and transformation of land through cultivation and urban expansion may break up the landscape and prevent fire spread. Removal of fuel for building, domestic cooking, and heating purposes, may also reduce fire spread (Saunders *et al.*, 1991; Frost, 1999).

All the factors illustrated in Fig. 1 vary spatially, and burnt area is unlikely to be controlled by the same combination of factors in different parts of the subcon-



**Fig. 1** Theoretical model of the factors affecting burnt area. Items in upper case are components of a fire regime, items underlined are direct drivers of fire. The rest (list of items on the left) are the indirect drivers of fire, which can be measured. Many of these factors covary: rainfall affects tree cover, population density is correlated with grazing density.

continent. This paper aims to determine which environmental and human drivers are important in affecting fire regimes in different areas of southern Africa; and specifically, to test whether human population densities have a positive or a negative effect on burnt area. Humans can affect fire regimes directly, by altering the ignition regime, and indirectly, by reducing fuels and fragmenting the landscape, thus reducing fuel continuity. Thus, increasing human densities could be predicted both to increase the incidence of fire, and decrease the extent of fire, and it is unclear which effect would be more important in determining total burnt area.

Similarly, areas with high rainfall have high fuel loads but also would have many perennial rivers, which might act as barriers to fire spread; and shorter dry seasons, which might limit the number of fires that could occur. We aim to find out whether, and under what circumstances, factors that decrease fire ignition and spread outweigh the positive effects of increased grass biomass.

Moreover, there are questions concerning the importance of fire management. Commercial farmers prepare fire breaks, actively suppress unplanned fires, and try to burn at certain times of year with set return periods (VanWilgen & Scholes, 1997). Fire management in communal land is more varied; fires are lit for a variety of reasons and it is not common to suppress actively burning fires (Mendelsohn, 2002; Verlinden & Laamanen, 2006). Early-season burning in communal areas is commonly undertaken, which could break up fuel loads for later fires and reduce total burnt areas (Frost, 1999; Laris, 2005). We aim to investigate whether land tenure differences have any apparent effect on burnt area.

Ultimately, we want to determine whether human or climatic influences are more important in driving fire regimes in southern Africa, so as to gain insights into how global change – both human and climate induced – will affect fire regimes. This study is based on only 1 year of data, and thus concerns itself largely with variation in space, but insights gained from this analysis will contribute towards predictive models of burnt area for the subregion.

## Data and methods

### Data

Spatially explicit burned area data and data on the environmental factors thought to affect burnt area for all of Africa south of the Equator (including Madagascar) were assembled for the 2003 fire season. Summary statistics of each dataset in the Lambert Azimuthal equal area projection were derived with respect to fixed, nonoverlapping square windows, defining the mean,

median, or percentage value within each window across the study area. A 100 km × 100 km window size was adopted as a result of window size sensitivity analysis, resulting in 899 sample points.

Selection of appropriate environmental factors was based on a conceptualisation of the direct and indirect drivers of fire in southern Africa (Fig. 1). Fuel load, fuel moisture, fuel-bed continuity, and wind speed all affect the spatial extent of an individual fire (VanWilgen & Scholes, 1997; Stott, 2000). These factors, together with ignition frequency, are direct drivers of annual burnt area and are influenced by a number of indirect drivers (Fig. 1). Eleven spatially explicit datasets describing the indirect drivers were used (illustrated on the left side of Fig. 1), of which four represented controls on fuel load, one on fuel moisture, three on fuel continuity, and three on ignition frequency. Wind speed and relative humidity at the time of burning are important in determining fire size and intensity (McArthur, 1966; Trollope, 1984) but were not included, as these data are not available for the entire study region.

The 11 independent variables are summarised in Table 1 and are described below, after description of the burned area data. Most of the independent variables are available as derived data products, and their sources and attributes are noted in Table 1. Others, including grazing density, land tenure, extent of transformed land, length of the dry season, topographic roughness, and lightning strikes, were derived from the best available spatial information.

### Annual burned area

Data for the main burning season (March–November) of 2003 were used in this study. The dependent variable, annual burned area, was derived from the NASA Moderate Spatial Resolution Spectroradiometer (MODIS) burned area product, that defines the 500 m location and approximate day of burning to a precision of 8 days (Roy *et al.*, 2005b; Roy *et al.*, 2008). The MODIS design includes features specifically for monitoring fires (Kaufman *et al.*, 1998). The MODIS burned area product maps the spatial extent of recent fires and is available as a monthly summary product. It is being generated on a global systematic basis in support of the global change community in conjunction with a number of other NASA-funded satellite products (Justice *et al.*, 2002). Southern Africa was selected as the first regional test for the burned area product that has been validated using independent reference data collected by the Southern Africa Fire Network (Roy *et al.*, 2005a) detecting about 85% of the true burned area. At the time of writing, only 1 year of burned area data was available. In addition, the 1 km MODIS active fire product that defines the

**Table 1** Spatial data on environmental drivers used as inputs to the regression tree model

Direct driver of burnt area	Indirect driver (input surface)	Name	Units	Summary statistic	Original resolution (at equator) (km)	Source and attributes	Notes
Fuel load	Mean annual rainfall (previous 2 years)	rainfall	mm yr <sup>-1</sup>	Mean	28	TRMM monthly 0.25° rainfall product 3B43 (V6) (Huffman <i>et al.</i> , 1995)	Rainfall positively affects grass production. This effect is cumulative over years. Other work (Balfour & Howison, 2002; Van Wilgen <i>et al.</i> , 2004) identified the previous 2 years of rainfall as the best predictor of total burnt area between years
	Soil fertility	sand	%	Mean	9	IGBP global soil data products 5 arc min (IGBP, 2000)	Soil fertility positively affects grass production. We used the percentage of sand in the soil as a very simple indication of the availability of nutrients and water to plants
	Tree cover	treecover	%	Mean	0.5	MODIS Vegetation Continuous Fields 500 m (Hansen <i>et al.</i> , 2003)	Tree cover has a negative impact on grass biomass (Scholes & Archer, 1997), especially above about 40% cover (Scholes, 2003)
	Grazing	grazing	kg km <sup>-2</sup>	Median	6	Global Livestock Distributions 3 arc min (FAO, 2005) filled using (Fritz & Duncan, 1994) and unpublished data (see text)	Grazer density affects how much of this fuel will be left in the system to feed fires. Grazing is likely to be correlated with soil fertility, as more fertile areas support more grazers (Fritz & Duncan, 1994).
Fuel moisture (flammability)	Length of the dry season	dryseason	Months	Mean	28	Calculated using the TRMM monthly 0.25° rainfall product 3B43 (V6) (Huffman <i>et al.</i> , 1995)	The longer the dry season, the drier the fuel, and the more time spent in a flammable state. Dry season length is defined as the inverse of the number of months it took for 70% of the rain to have fallen (i.e. 12-length of wet season = length of the dry season)

*Continued*

Fuel continuity	Topographic roughness	topography	m	Mean	0.09	Calculated using SRTM 90 m digital topographic data Version 2 (STRM, 2005)	This gives an indication of the flatness of the terrain, and how many natural barriers there are to fire spread. It is defined as the standard deviation of a $3 \times 3$ block of a 90 m resolution DEM (see Russell-Smith <i>et al.</i> , 2007)
	Road density	kmroads	km	Sum	N/A	MapStudio (www.mapstudio.co.za) vector road data for Africa, 2000.	This is a measure of human impacts on fuel continuity.
	Fraction of transformed land	transformed	%	%	1	Global land cover 2000 1 km (Mayaux <i>et al.</i> , 2004)	This is a measure of human impacts on fuel continuity – Land classified as cultivated or settlements by the Global Land Cover were considered transformed. The fraction of transformed land in the analysis window was the input variable
Ignition frequency	Mean lightning strikes over the burn season.	lightning	#hits $\text{km}^{-2}$	Mean	55	GHRC LIS/OTD 0.5 Degree High Resolution Full Climatology (GHRC, 2003)	Areas with more lightning strikes might have higher lightning-induced ignitions. Only lightning that occurs at the same time as fire could be a source of ignitions (see Fig. 2). Therefore, the months during which 90% of the burnt area occurred were identified, and the average lightning frequency over these months was calculated
	Population density	popdensity	# people $\text{km}^{-2}$	Median	5	Ciesin Gridded Population of the World 2.5 arc min Version 3. Adjusted estimates for 2000 (Ciesin, 2005).	It has often been suggested, and one study has shown (Keeley <i>et al.</i> ), that increased population densities result in increased numbers of fires
	Percentage of communal land	communal	%	%	N/A	Data surface created using individual country data (see text for details)	We hypothesised that ignitions might be more frequent on communally managed land where fire is used freely (Mendelsohn, 2002) compared with land managed for commercial or conservation purposes – which often have strictly controlled fire schedules

TRMM, Tropical Rainfall Measuring Mission, GHRC, Global Hydrology Resource Centre, MODIS, Moderate Spatial Resolution Spectroradiometer.

location of fires observed at the time of satellite overpass (Giglio *et al.*, 2003) was used to study ignition incidence. Interpreting polar-orbiting satellite active fire data is complicated by the fact that only fires actively burning at the time of satellite overpass and under cloud free conditions are detected, but when interpreted with this caveat in mind, the active fire product provides qualitatively different information from the burnt area product (see 'Results'). Further information and MODIS fire product examples can be found at <http://modis-fire.umd.edu/MCD45A1.asp>.

#### *Grazing*

The grazing intensity data (kg of live mass km<sup>-2</sup>) were created by summing cattle, sheep, and goat biomass per unit area from the FAO livestock distribution dataset (FAO, 2005). The FAO data do not cover protected areas so the effect of indigenous large mammal grazers was estimated for these regions using wildlife census data from Kruger, Gorongosa, Etosha, and Serengeti National Parks, as well as published carrying capacity information (Fritz & Duncan, 1994).

#### *Land management*

Land tenure was used as an easily derived spatial indication of different land management practices in the region. Three different land management (land tenure) classifications were identified: communal, commercial, and protected areas. The surface was created by combining available land tenure maps for South Africa, Namibia, Botswana, and Zimbabwe with the World Protected Areas map (UNEP-WCMC, 2006). Each map was interpreted to identify communal, commercial, and protected areas. We assumed that the extent of commercial rangeland in Angola, Zambia, Zaire, Mozambique, and Tanzania was negligible, and that the major land tenure types in those countries were communal and protected areas.

#### *Land transformation*

The Global Land Cover 2000 Africa product (Mayaux *et al.*, 2004) was simplified into a binary map of 'transformed land' (crops, plantations, or urban use) and 'untransformed land' (all other categories). This was used as a measure of the extent to which human activities have fragmented the fuel load.

#### *Accumulated rainfall*

The monthly Tropical Rainfall Measuring Mission (TRMM) best-estimate precipitation rate product (Huff-

man *et al.*, 1995) was summed between July 2001 and June 2003 and divided by 2 to give the preceding 2-year average rainfall (see Van Wilgen *et al.* (2004) for justification of this metric).

#### *Length of the dry season*

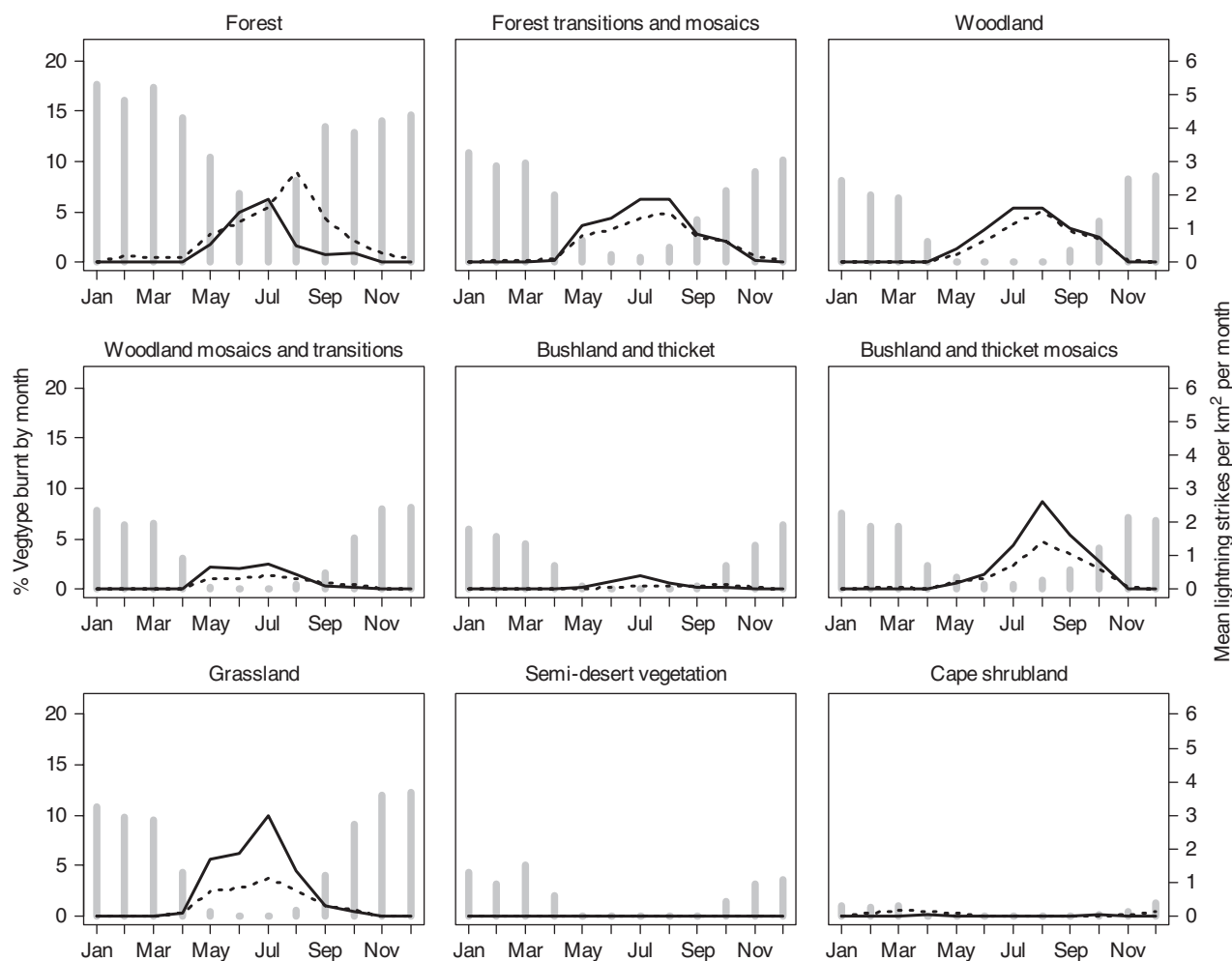
The dry season length (months) was derived using the monthly TRMM best-estimate precipitation rate product (Huffman *et al.*, 1995). For each TRMM pixel, 12 monthly rainfall estimates from July 2002 to June 2003 were ranked and summed in descending order until at least 70% of the annual rainfall was reached. The remaining number of months (representing <30% of the annual rainfall) was considered as the length of the dry season for that pixel. These values ranged from 3 months in the high-rainfall belt near the equator, to 9 months in the subtropical arid zones in the continental interior. This metric represents an improvement on dry season metrics developed in the literature (Spessa *et al.*, 2005; Russell-Smith *et al.*, 2007) in that it is independent of total annual rainfall or an *ad hoc* identification of the wet season. These are strongly seasonal climate systems, and a sensitivity analysis showed that the data were not sensitive to the choice of the 70% cut-off.

#### *Topography*

Topographic roughness was derived from Shuttle Radar Topography Mission (SRTM) elevation data (STRM, 2005) as the standard deviation of the SRTM elevation values in 3 × 3 90 m pixel windows (following Russell-Smith *et al.* (2007) and resampled to 500 m using the nearest neighbour method). At the scale of this analysis (values summarised over 100 km windows), topographic effects are likely to be manifest in terms of the negative effect that a highly dissected landscape has on fire spread (Frost, 1999; Russell-Smith *et al.*, 2007). At finer scales, the topographic position of individual fires affects the rate of spread (Viegas, 2004) and the fuel moisture (Heyerdahl *et al.*, 2001), but we expect that these effects would not be apparent in our analysis.

#### *Lightning frequency*

Most lightning in these systems occurs in the rainy summer months when there is very little fire (Fig. 2). To extract a metric for lightning ignitions we included only the lightning strikes that could have been ignition sources for the fires in the system. Data from the Global Hydrology Resource Centre (GHRC, 2003) were used to compute the mean frequency of lightning strikes per km<sup>2</sup> during the fire season (defined as those months in which 90% of the fires occurred). No correction was



**Fig. 2** Seasonal distribution of lightning strikes and burnt area in the 12 major vegetation types identified by White (1983). Grey bars show mean monthly lightning strike frequency; solid lines show the percentage of the area burnt defined by the Moderate Spatial Resolution Spectroradiometer (MODIS) burned area product (Roy *et al.*, 2005b); dotted lines show the percentage of fire-affected area defined by the 1 km MODIS active fire product (Giglio *et al.*, 2003).

made for cloud-to-cloud strikes because, lacking any data to the contrary, it was assumed that the proportion of these strikes was relatively constant across the study region.

#### Data preprocessing

Considerable effort was taken to ensure that these data were precisely coregistered, and projected in a way that preserved the data integrity. All the data were reprojected into the same Lambert Azimuthal equal area projection with an African projection (centre of latitude, 25°, centre of longitude, 15°; sphere radius 637 0997 m). The raster data products were reprojected with nearest neighbour resampling to maintain the pixel values, and resampled with 1 km or 500 m output pixel dimensions to reduce nearest-neighbour resampling pixel shifts (i.e. position errors) (Dikshit & Roy, 1996). Similarly, the

vector data were converted into raster thematic layers with 1 km output pixel dimensions.

Summary statistics of each dataset in the Lambert Azimuthal equal area projection were derived with respect to fixed, nonoverlapping square windows, defining the mean, median or percentage value within each window across the study area. A 100 km × 100 km window size was adopted as a result of window size sensitivity analysis. Road density (km) was derived by summing the length of roads in the 100 km × 100 km window. Mean values were used for the other independent datasets (Table 1), except for grazing and human population density. Median values were used to summarise these latter data to reduce the skewness introduced by high human populations confined to relatively small urban areas, and similarly to reduce the influence of some unrealistic outliers in the FAO cattle data (W. Wint, personal communication, 2006). The annual

percentage burned was found by summing within each window the number of 500 m pixels that were labelled as burned in the period March–November 2003. The proportion of nonland (i.e. coastal and inland water bodies) and invalid (i.e. missing, cloud contaminated, unavailable, unmapped) data within each window were computed. Windows for which any of the datasets listed in Table 1 had greater than one-third nonland and invalid pixels were excluded from the analysis.

### Analysis

Conventional statistical models such as General Linear Models are inappropriate for investigating the drivers of burnt area in the context of this study because many of the relationships are likely to be nonlinear, with non-additive predictor–response interactions, and several of the independent variables are highly correlated (see Table 2). Decision trees (regression trees) are hierarchical classifiers that predict class membership by recursively partitioning data into more homogeneous subsets, referred to as nodes (Breiman *et al.*, 1984). They accommodate abrupt and nonmonotonic relationships between the independent and dependent variables and make no assumptions concerning the statistical distribution of the data. When run using continuous data, a sum of squares criterion is used to split the data into successively less varying subsets. This splitting procedure is followed until a perfect tree is created or until preset conditions are met for terminating the tree's growth. It is then possible to identify the variables and the split conditions that result in the final prediction. For our purposes, it is this latter property that make trees particularly advantageous; the tree structure enables interpretation of the explanatory nature of the independent variables.

A random forest regression tree procedure was used. Like other bootstrapping procedures, random forests improve the predictive ability of regression tree models and reduce overfitting (Breiman, 2001; Prasad *et al.*, 2006). A

large number of regression trees are grown, each time using a different random subset of predictor variables, and keeping a certain percentage of data aside ('reserved'). In our case, we grew 1000 trees, using 3 of the 11 variables at each split, and using 66% of the data each time. The default minimum node size of five (Breiman, 2001) was used, meaning no node with fewer than five cases was split. Predictions were run on the reserved data each time a tree was grown, and the final prediction for each data point was the mean of the predicted values. The analysis was undertaken using R software (<http://www.r-project.org/>), using the RANDOMFOREST package.

### Importance of independent environmental variables

There are several metrics for determining how important different variables are to the final prediction (Breiman, 2001). We used a method which takes the difference in mean square error (MSE) between a test sample, and the test sample when that variable is randomly permuted, calculated using the reserve data for each tree, for each variable (Breiman, 2001). The differences are averaged over all 1000 trees, and then normalised by the standard error. This value, therefore, provides a measure of how much the predictive ability of the model is reduced when the effect of a certain variable is excluded. When looking more specifically at the effect of individual drivers on burnt area, a piecewise quantile linear regression method (Koenker, 2005; Sankaran *et al.*, 2005) was used to fit the upper boundary of burnt area and to identify breakpoints. This was done using QUANTREG in R statistical software. The 99th quantile was used to identify the upper limits of burnt area under different drivers.

### Split conditions

In the random forest procedure many different trees are grown and the average result taken, so one cannot

**Table 2** Correlations between the 11 independent environmental variables used in the analysis

treecover										
<b>0.75</b>	rainfall									
<b>-0.50</b>	<b>-0.53</b>	dryseason								
-0.25	-0.08	0.00	grazing							
0.05	0.17	-0.15	0.48	popdensity						
-0.38	-0.29	-0.07	0.28	0.13	kmroads					
0.28	0.41	-0.40	0.07	0.32	0.09	topography				
-0.04	-0.01	0.02	<b>0.50</b>	0.35	0.27	0.08	transformed			
0.47	<b>0.54</b>	-0.04	-0.03	0.22	<b>-0.51</b>	0.11	-0.02	communal		
<b>0.62</b>	0.35	<b>-0.55</b>	-0.15	-0.02	-0.11	0.17	-0.05	0.06	lightning	
-0.23	-0.22	0.10	-0.19	-0.26	0.00	-0.32	-0.20	-0.14	-0.10	sand

Bold values represent significant correlations ( $P < 0.05$ ).



explore the explicit split conditions for classifying data into different groups. As the particular conditions resulting in high and low burnt areas were of interest to us, we took this approach a step further and ran one more regression tree using the predicted data from the random forest permutation as an input. We pruned this tree using a complexity parameter of 0.01 (see R documentation 'rpart.control' for an explanation of this parameter). This tree was stable and explained 85% of the variance in the input data which was significantly ( $P < 0.001$ ) more than a random tree [see Rejwan *et al.* (1999) for significance-testing methods for regression trees]. The original observed burnt area data were then classified using this regression tree, and the splits and nodes analysed to determine the combinations of environmental conditions that result in high and low burned area proportions across the subcontinent.

### Spatial autocorrelation

Spatial autocorrelation in the burned area data (Lynch *et al.*, 2006) and the predictor variables may result from various causal mechanisms (e.g. physical and biological processes), acting both simultaneously and additively, and from variables and processes not quantified in this study. The scale of the analysis and the summary units used to aggregate the data into  $100 \text{ km} \times 100 \text{ km}$  windows also imposes a spatial structure on the data. It is established that autocorrelated data violate the assumption of independence of many statistical procedures (Cliff & Ord, 1981; Legendre & Legendre, 1998). Approaches to resolve this issue have been to manipulate the sampling scheme to avoid autocorrelated observations or to attempt explicitly to incorporate spatial dependence into the model.

To date, no technique to incorporate spatial dependence into decision trees has been reliably demonstrated, and this remains an area of active research (Miller *et al.*, 2007). In this analysis rainfall, tree cover, and topography are all explicitly included as predictor variables that account for the major geographical gradients. We also subsample the data in the random forest procedure (using 66% of the data each time) that will further reduce the spatial dependency of the data. However, we admit to a degree of uncertainty related to the statistical tests that may result due to inclusion of spurious explanatory variables (Lennon, 2000); for these reasons only, the most important splits in the regression tree results are considered in 'Results'.

### Window size sensitivity analysis

Geographical analyses of this kind are sensitive to the scale of the analysis; both with respect to the size and

location of the windows and to the nature of the summary units used to aggregate the data (Openshaw, 1984; Unwin, 1996). Rather than select an arbitrary window size, a sensitivity analysis was undertaken to evaluate an appropriate size for the analysis. This was undertaken by applying the random forest regression tree procedure to the data several times, each time using a progressively larger window size and recalculating the summary statistics (fraction, percentage, mean or median). Window side dimensions varying from  $30 \text{ km} \times 30 \text{ km}$ , approximately equivalent to the coarsest available input data ( $0.5^\circ$  lightning data, Table 1), to an order of magnitude greater,  $300 \text{ km} \times 300 \text{ km}$ , were considered, providing 10876 to 114 windows across the study area, respectively.

## Results

### Window size analysis

The number of nodes in the final tree declined steeply with increasing window size, and the predictive ability ( $r^2$ ) of the random forest model declined less steeply (Fig. 3). The  $r^2$ -values are calculated for each window size from the linear relationship between the burnt area values predicted by the statistical model and the observed percent burnt area. Window dimensions between  $100 \text{ km} \times 100 \text{ km}$  and  $150 \text{ km} \times 150 \text{ km}$  show some variability in  $r^2$ , perhaps related to the original

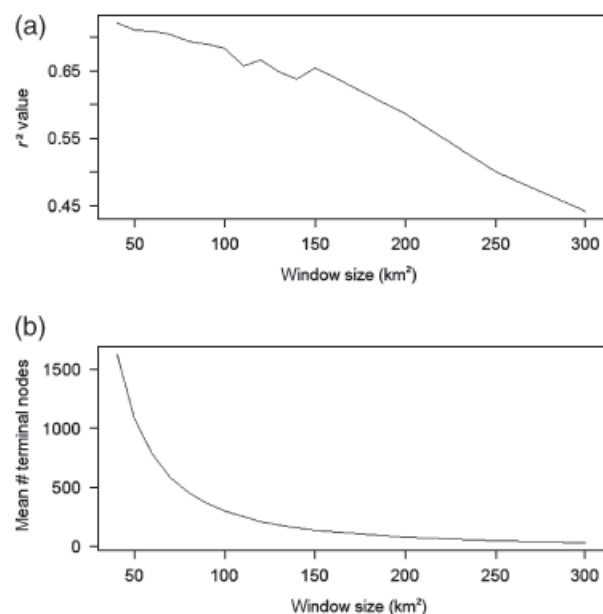


Fig. 3 Sensitivity analysis of the effect of increasing window size on (A) the predictive ability ( $r^2$  of predicted vs. observed burnt area) and (B) the number of terminal nodes, in the random forest regression tree model.

resolution of some of the input data relative to the window sizes and locations. The 100 km × 100 km window was chosen as an acceptable trade-off between overfitting the model (too many split conditions), and reducing the predictive ability (by averaging the data too much).

#### *Spatial patterns of burnt area*

The MODIS 500 m burned area product identified 18% of the observable study area as burnt in 2003. Some 26% of the region could not be classified as burned or unburnt due to persistent cloud and missing MODIS observations (Roy *et al.*, 2005a), so the actual burnt area statistic for Africa south of the equator could range from 13% to 39% (from all unclassified pixels unburnt or all burnt, respectively). The majority of the unclassified pixels occurred over the equatorial parts of Gabon, Congo, and the Democratic Republic of Congo (DRC) and are associated with persistent cloud at the time of MODIS overpass. The DRC and Angola had the highest burnt areas: 56% and 41%, respectively, of the valid pixels in these countries were burnt in 2003. Although we were unable to determine the state of burning in the invalid pixels, it is likely that the high burn percentages are representative of the country totals (B. Muhigwa, personal communication). In contrast, 2.5% and 1.7% of South Africa and Namibia, respectively, burnt. These results are consistent with other burned area estimates for the region: total burned area derived from the SPOT vegetation product for 2000 for the same area was calculated at 17% (Silva *et al.*, 2003), and the spatial distribution of burning reflects that of independently detected MODIS active fires (Roy *et al.*, 2005a). The annual fraction burned data are shown in Fig. 5a, summarised for the 100 km × 100 km window size adopted for the analysis.

#### *Relative importance of environmental variables*

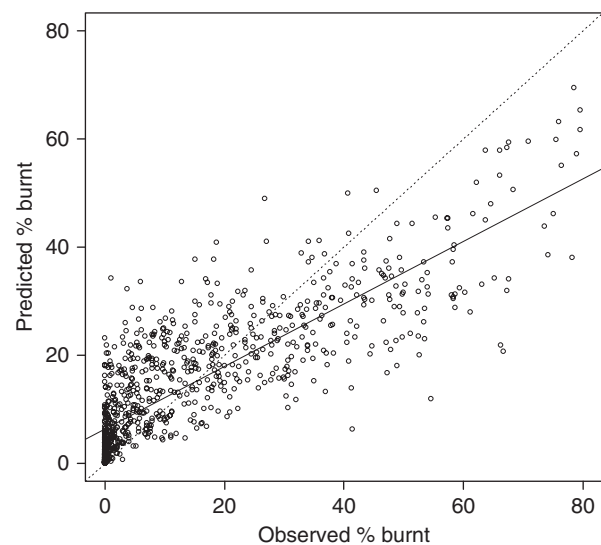
A linear regression between predicted values from the random forest and observed burnt areas had an  $r^2$  of 0.68 (Fig. 4), and the model captures the spatial distribution of burning well (Fig. 5). The model underpredicts the high burnt areas and overpredicts the low burnt areas, as would be expected from a procedure that groups data into homogeneous classes.

The analysis identified tree cover, rainfall, and dry season length as the most important predictors of burnt area across the study region (Fig. 6). Human activity also appears to influence burnt area, with grazing, population density, and road density all identified as moderately important – increasing the MSE by between 35% and 40% (see 'Data and methods' for definition of

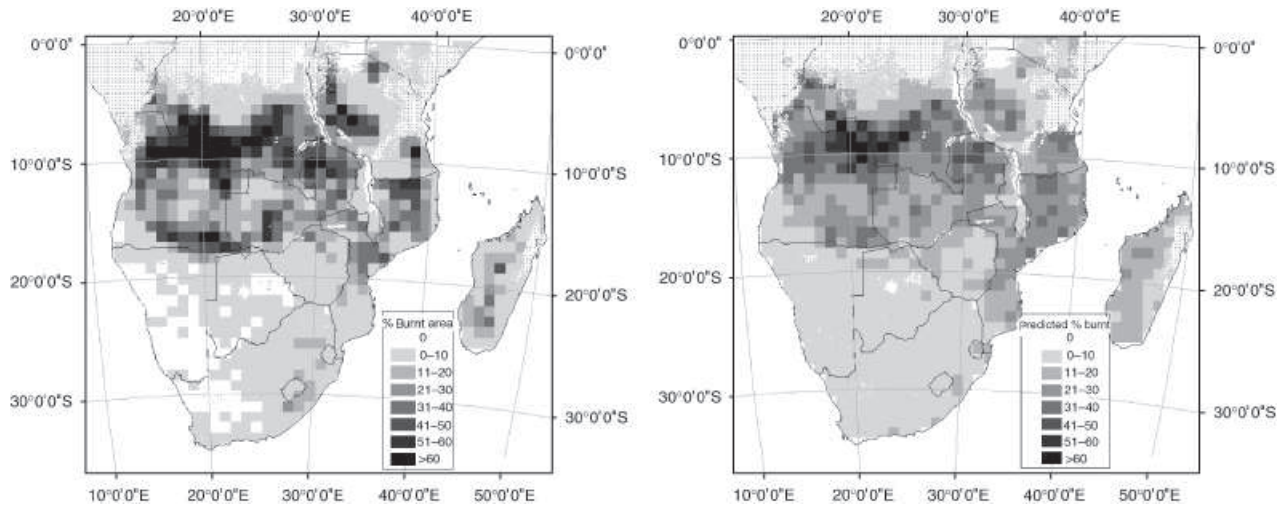
importance). Sand percentage – an inverse measure of soil fertility – was the least important variable, probably because grazing density is a better indicator of the effects of soil fertility. Lightning density was also less important. As can be seen from Fig. 2, fires occur at very different times of year from the peak density of lightning strikes in all vegetation types. This supports commonly held perceptions that humans are the main sources of ignition in Africa (Frost, 1999; Sheuyange *et al.*, 2005). However, it does not totally exclude lightning as an important source of ignition at certain times of year (such as the early wet season), or in places where there are few people (Edwards, 1984).

#### *Split conditions*

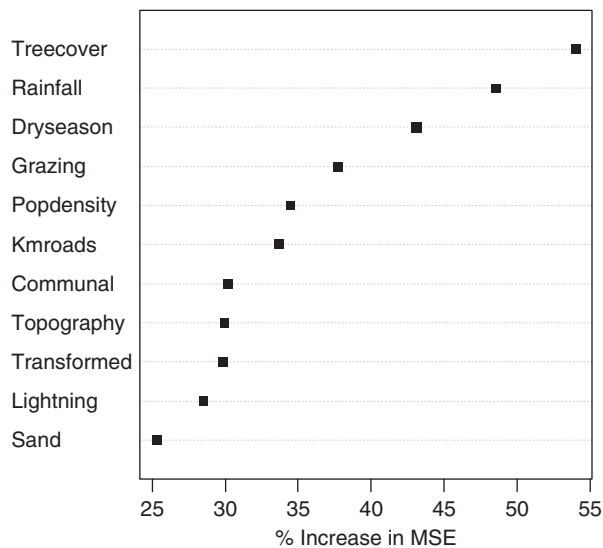
The conditions that result in high or low burnt areas on the continent were identified by cascading the observed burnt data through the final regression tree created from the random forest predictions (Table 3). The reliability of this prediction is slightly decreased from the original random forest predictions, ( $r^2 = 0.57$ ,  $P < 0.001$ ) but because the split conditions are available, these results give an indication of the mechanisms driving



**Fig. 4** Predicted burnt area calculated from a random forest regression tree run on 2003 burnt area data using 11 independent variables (Table 1). The dotted line represents an exact 1:1 relationship, the solid line shows a linear regression of these data. The explained variance ( $r^2$  of regression line) is 68% but the model overpredicts at low burnt areas and underpredicts at high burnt areas. One thousand trees were run, excluding 33% of the data each run ('reserve' sample). The predicted value is calculated as the mean value predicted each time a data point was included in the reserve sample (see 'Data and methods').



**Fig. 5** Observed (left) and predicted (right) percent burnt areas for 100 km  $\times$  100 km windows across southern Africa. Predictions are mean values for the reserve samples in the random forest (see 'Data and methods'). Dark stipples represent areas where cloud or missing Moderate Spatial Resolution Spectroradiometer (MODIS) data preclude burned area mapping for more than 5 months of the year.



**Fig. 6** Relative importance of each independent variable in determining burnt area for Africa south of the equator. Importance is measured by testing how the accuracy of the results is affected if the input variable is randomly permuted. It is calculated as the increase in mean squared error (MSE) that occurs in the predicted values when each variable is randomly permuted in the test set. Importance is calculated for each variable, for each tree, then averaged, and normalised by the standard error. See Table 1 for a definition of variables.

fire on the continent, and how these drivers change over space.

Two different sets of conditions can result in very low burnt areas (nodes 1 and 5; Table 3). Areas with low tree cover have a mean predicted burnt area of less than 1%. These areas coincide with areas of very low rainfall

(mean rainfall 288 mm), and it is likely that low rainfall, rather than low tree cover, is the cause of this. The low rainfall does not permit the accumulation of sufficient fuel to sustain extensive fires. The second condition resulting in low burnt area is that the length of the dry season is less than 6 months. Figure 7a shows the parts of Africa where these two sets of conditions hold true; covering approximately 31% of the study area.

Extremely low burnt areas, therefore, seem to be correlated with a certain set of climatic factors. However, human activities can *decrease* burnt areas in circumstances that otherwise would result in intermediate to high burnt areas. For example, when tree covers are between 5% and 21%, a large range in burnt areas is possible (3–35%). The percentage of the landscape that actually burns in these areas depends on grazing density, and the density of the road network (nodes 2, 3, and 4; Table 3). Fire regimes in these parts of Africa, which cover around 28% of the total area (Fig. 7b), are therefore highly spatially variable, and can be modified by human activities.

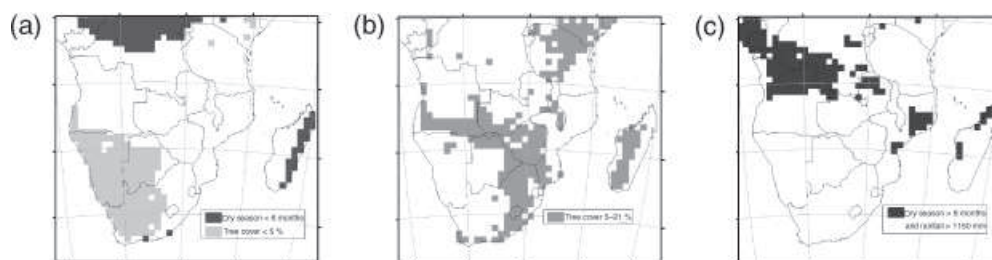
Where annual rainfall exceeds 1150 mm a year and the dry season is longer than 6 months (node 9), annual burnt area exceeds 20%. Although high population densities, variable terrain, and high tree covers can reduce burnt area, regular surface fires are an inevitable consequence of climate in these parts, which cover 17% of the subcontinent (Fig. 7c).

The ability of the regression tree to predict burnt area was more reliable at the extremities, where either a relatively large or small proportion of the area would be predicted to burn (Fig. 8). Under conditions that led to intermediate burnt areas (between 20% and 40%)

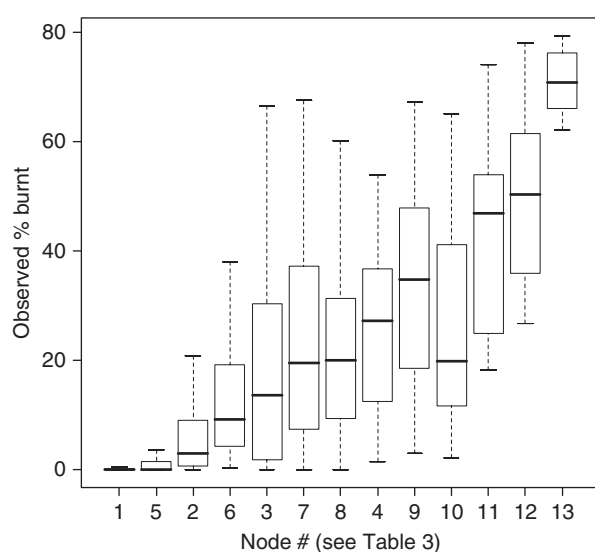
**Table 3** Split conditions identified by a regression tree run on the random forest predictions

Split conditions	Mean percent burnt	Node #
treecover < 21	0	1
treecover < 5		
treecover ≥ 5 kmroads ≥ 355	3	2
Kmroads < 355 grazing ≥ 5	14	3
grazing < 5	35	4
treecover ≥ 21		
dryseason < 6	0	5
dryseason ≥ 6 grazing ≥ 6 kmroads ≥ 333	9	6
kmroads < 333	20	7
grazing < 6 rainfall < 1150	20	8
rainfall ≥ 1150 popdensity ≥ 15	27	9
popdensity < 15 treecover ≥ 38	20	10
treecover < 38 topography ≥ 0.4	47	11
topography < 0.4 rainfall < 1300	50	12
rainfall ≥ 1300	71	13

Observed burnt area data were classified with this regression tree, to produce the mean % burnt values.



**Fig. 7** Areas where the fire regime is under varying degrees of climatic control. (a) Areas of very low tree cover (and rainfall) always have low burnt areas (grey), as do areas with short dry seasons (black). (b) Areas where the burnt area is most under human control: depending on population density, grazing, road density, and the percentage of transformed land these areas can range from 3% to 35% burnt. (c) Areas where significant percentages of the landscape burn, and human activities reduce, but cannot eliminate, fire. Here dry season is measured as the number of months with less than 30% of the annual rainfall. These maps were created by plotting spatially the areas in Africa that fulfilled particular environmental conditions identified in Table 3.



**Fig. 8** Median (horizontal thick line) and range ( $\pm 25\%$  quantiles) of burned area class values identified by the regression tree (Table 3). Some burnt area classes have less variance than others, particularly those that predict the smallest and the largest burnt area. The regression tree was created using the predicted values from a random forest bagging procedure, and the 11 environmental inputs in Table 1. This resulted in a stable tree, which was then used to classify the observed burnt area values.

predictions were less certain, probably because there are a number of different combinations of factors that can result in this range of burnt areas, and these are not fully captured by the first 12 splits of the regression model.

#### *The effect of people*

Population density always had a negative effect on burnt area in the regression tree model. Fig. 9a, and a

piecewise quantile linear regression model (Koenker, 2005; 99th percent quantile, run using QUANTREG in R) indicate that the negative effect of people on burnt area holds for population densities greater than 10 people  $\text{km}^{-2}$ . Below this population density, the slope of the quantile regression is not significantly different from 0.

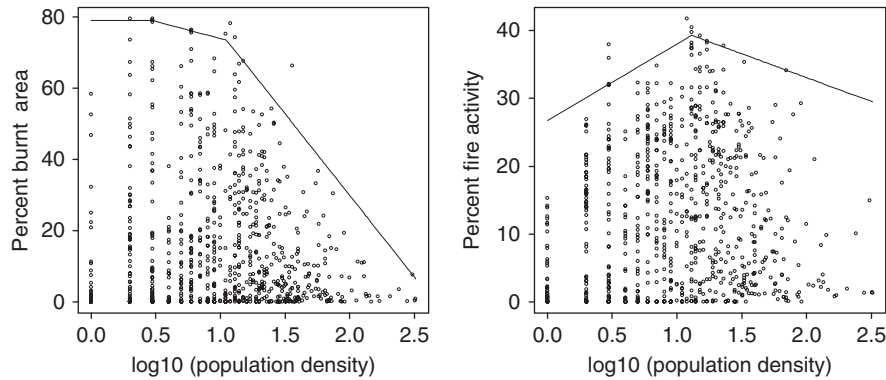
The relationship between population density and *number* (rather than *area*) of fires appears to be quite different however (Fig. 9b). The results suggest that increasing human population densities up to around 10 people  $\text{km}^{-2}$  are associated with more fires, but that densities higher than 10 people  $\text{km}^{-2}$  are associated with fewer fires (Fig. 9b). The difference between Fig. 9a and Fig. 9b implies that as population density decreases, the size of individual fires must increase, we suggest due to less fragmented fuel beds.

## Discussion

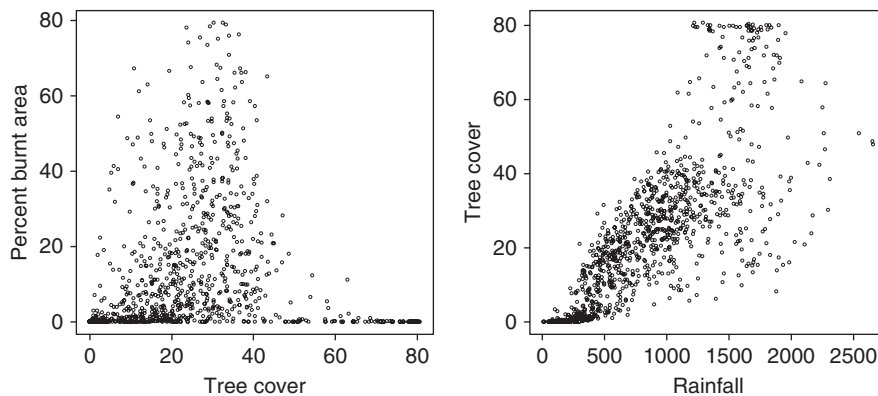
### *Drivers of fire: fuel and weather*

In order for a fire to start and spread three conditions must be met: there must be an ignition event, there must be flammable fuel, and the weather conditions must be suitable (VanWilgen & Scholes, 1997; Stott, 2000). Moreover, the fuel bed must be adequately continuous, or the fire will not spread. At a subcontinental scale, our analysis suggests that annual burnt area is controlled mainly by the second condition: the fuel. The four most important variables identified by the random forest procedure (Fig. 6) either affected fuel loads (rainfall, tree cover, grazing) or fuel moisture (length of the dry season). The continuity of the fuel bed was also shown to be important, as high road densities and variable topography both limit burnt area.

We were unable to include weather conditions in our statistical model – the daily data necessary to produce daily fire danger indices were not available at the scale



**Fig. 9** Burned area (percentage of 100 km × 100 km window) plotted against log population for the Moderate Spatial Resolution Spectroradiometer (MODIS) burned area product (left) showing the negative effect of people on burnt area at densities over about 10 people km<sup>-2</sup>; and for the MODIS active fire product (right) showing a different pattern, where decreasing densities of people appear to decrease fire activity. A piecewise quantile linear regression (Koenker, 2005; 99th percent quantile, run using QUANTREG in R) of these data is superimposed.



**Fig. 10** Burned area (percentage of 100 km × 100 km window) plotted against (a) rainfall and (b) percent tree cover. The maximum possible percent burnt area declines rapidly at tree covers greater than 40%.

of Africa south of the equator. However, it has been shown in several fire-prone systems that a small number of large fires account for the majority of the area burnt (Yates & Russell-Smith, 2002; Dickson *et al.*, 2006). These very large fires are thought to be caused by extreme weather conditions (McArthur, 1968; Bessie & Johnson, 1995; Moritz, 2003), in particular, periods of several days of hot, dry winds following a good rainy season. Including daily-resolution fire weather would undoubtedly improve the predictive power of the model.

Where tree cover exceeds 40%, the maximum possible percent burnt area declines rapidly (Fig. 10a). This threshold presumably results from a reduction in grass fuels as tree density increases, and was also identified by the regression tree procedure (Table 3 node 10). There is much evidence for a nonlinear interaction

between fire and vegetation structure in these systems – and forest and fire-maintained grasslands are identified as alternate stable states (Bond *et al.*, 2005). If the sudden drop in burnt area that we found is valid, and not an artefact of the remotely sensed data used, then it appears that 40% tree cover is the threshold at which such a system switch might occur, fires might be kept out, and a tendency to canopy closure would proceed. Empirical data supports this – identifying a nonlinear response of grass productivity to tree cover with an inflection point around 35% and 40% cover (Scholes, 2003), and a switch in the understorey species composition to more forbs (Malaise, 1978).

Our data show that tree cover can only exceed 40% in systems with more than about 800 mm of rainfall (Fig. 10b). Thus, these results corroborate Sankaran *et al.* (2005), who found that a rainfall threshold of 784 mm

was the dividing point between stable (rainfall-maintained) and unstable (fire/disturbance-maintained) savannas.

#### *Drivers of fire: ignitions*

We found no indication that the frequency of ignition limited or promoted burnt area in the region. Lightning frequency over the fire season came out as one of the least important predictors, as did land tenure (there is a perception, unsupported by our data, that fires are more commonly ignited on communal land than on privately held land). Moreover, low human population densities, which should be associated with fewer ignitions (Kealey *et al.*, 1999; Stott, 2000), do not appear to reduce burnt area (Fig. 9a). The opposite was true; high human densities ( $>10$  people  $\text{km}^{-2}$ ) resulted in less area being burnt, probably because of the effect that people have in fragmenting the landscape (Saunders *et al.*, 1991) through cultivation, grazing livestock, fuel-wood collection, roads, or possibly by suppressing fires.

Theoretically, there must be a point at which a lack of ignition sources should limit the total area burnt, but our analysis did not reveal any set of circumstances where this might be the case. Low population densities were associated with a reduction in the number of fires identified by the MODIS active fire product but there was no concomitant reduction in burnt area (Fig. 9a and b). These results have two implications. Firstly, in sparsely populated regions, reduced ignition occurrence seems to be adequately compensated for by less-fragmented fuel beds, resulting in more extensive individual fires. Secondly, it takes very few people to provide sufficient ignition opportunities, and lightning can provide sufficient sources of ignition in areas where even this minimum population density is not reached. It is not possible to test this hypothesis without data on individual fire size, but Frost (1999) indicates that increased human densities are associated with smaller fires, and that the proportion of lightning-caused fires increases in drier, less-populated areas.

#### *Potential for change: where do people matter?*

The analysis makes some clear predictions on the conditions under which people can affect annual burnt area, and the conditions where climatic factors overwhelm the effect of people. Once the very dry ( $<300$  mm) and very wet ( $>1000$  mm) areas have been excluded, there remain large parts of Africa (Fig. 7b) where the burned area is responsive to influence by nonclimatic factors. In these regions, the burnt area fraction can range from 3% to 35% of the total area. Running a separate random forest procedure on this

data identified road density, grazing, fraction of transformed land (cultivated/urban), and population density as the four most important predictors of burnt area.

Given that major demographic and economic changes are predicted in 21st century Africa (Kirk, 1999; UNEP, 2002), it is likely that fire regimes, particularly in this intermediate-rainfall area, will also change.

#### *Potential for change: climatic limits to fire*

Importantly, this research also shows that there are large parts of southern Africa where human activities have little effect on fire regimes – and climatic factors either limit, or promote widespread fires. The analysis identified a tree density of  $<5\%$  (corresponding to a mean rainfall of around 288 mm) as a threshold below which very little fire activity occurs. The assumption that it is rainfall, not tree cover, that is causing the reduced burnt area is corroborated by field data on fire spread. Annual rainfall of 288 mm produces grass fuels of around 1000–1500  $\text{kg ha}^{-1}$  (Scholes, 2003), and Trollope's *et al.* (2002) fire behaviour equations for southern African savannas estimate a zero rate of spread at just under 1000  $\text{kg ha}^{-1}$ .

The other condition resulting in low burnt areas was a dry season of less than 6 months (i.e. for more than half of the year, monthly rainfall is significantly contributing to the total annual precipitation). Stott (2000) suggests that dry seasons of as little as 2.5 months are sufficient to provide dry grass fuels for burning. The fact that this analysis identifies 6 months of dry weather implies that it is not only the presence of dry, flammable fuels, but the proportion of the time that these flammable fuels are available that is important for determining burnt area. It also hints at a potentially interesting interaction between length of the dry season and ignition probability that should be further investigated. Ignitions might prove to be limiting in areas with very short dry seasons.

About 17% of the study region has the high rainfall and long dry seasons necessary for widespread fire (Fig. 7c). This relatively small area accounted for 37% of the area burnt in 2003. In these regions (core Miombo woodlands, covering northern Angola, southern Congo, and Northern Mozambique), human activities were still shown to reduce burnt areas, but never to very low levels (Table 3). These parts of Africa are the focus of current projects aiming to store carbon by reducing fire (peaceparks annual review: <http://www.peaceparks.org/>) so it is important to know the extent to which fire can be managed in these systems. Our results suggest that climatic and environmental conditions are likely to override human attempts to prevent fire – unless tree covers can be increased to 40% or more.

*Potential for change: climatic variability*

Much of southern Africa is characterised by high inter-annual variability in rainfall. Because all the other factors influencing burnt area on the continent change slowly, at decadal time scales, interannual variability in burnt area is likely to be driven largely by variation in rainfall and dry season length. Published analyses of long-term fire datasets support this hypothesis (Balfour & Howison, 2002; Van Wilgen *et al.*, 2004; Mulqueeney, 2005; Russell-Smith *et al.*, 2007). Thus, the limits of the fire-prone and nonfire-prone systems shown in Fig. 7 are strongly dependent on the rainfall preceding the year of study (2003), and could change quite dramatically during periods of extremely high or low rainfall. As more years of burnt area data become available, it will be possible to test whether the relationships between rainfall and burnt area identified at local sites also hold true over large areas, and apply these results to more mechanistic models of burnt area, including structured equation modelling. This, together with improved weather inputs, will be the next stage of the research.

**Acknowledgements**

We thank Dr John Mendelsohn, Dr Pauline Dube, Kolethi Gumbo, and Heidi van Deventer for providing the land tenure data, Minnie Wong for help in obtaining the MODIS active fire product, Drs Jeremy Russell-Smith and Mike Wimberly for their useful comments. The research was completed with funds from the CSIR SRP Global Change project, the EU CarboAfrica project, and NASA grant NNG04HZ18C.

**References**

Balfour DA, Howison OE (2002) Spatial and temporal variation in a mesic savanna fire regime: responses to variation in annual rainfall. *African Journal of Range and Forage Science*, **19**, 45–53.

Bessie WC, Johnson EA (1995) The relative importance of fuels and weather on fire behaviour in subalpine forests. *Ecology*, **76**, 747–762.

Boko M, Niang I, Nyong A *et al.* (2007) Africa. In: *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE), pp. 433–467. Cambridge University Press, Cambridge, UK.

Bond WJ, Van Wilgen BW (1996) *Fire and Plants*. Chapman & Hall, London, UK.

Bond WJ, Woodward FI, Midgley GF (2005) The global distribution of ecosystems in a world without fire. *New Phytologist*, **165**, 525–538.

Breiman L (2001) Random forests. *Machine Learning*, **45**, 5–32.

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) *Classification and Regression Trees*. Wadsworth & Brooks/Cole, Monterey, CA, USA.

Byram GM (1959) Combustion of forest fuels. In: *Forest Fire: Control and Use* (ed. Davis KP), pp. 61–89. McGraw Hill Book Co., New York, NY, USA.

Ciesin (2005) Gridded population of the world version 3 (GPWv3). [http://www.ciesin.org/download\\_data.html](http://www.ciesin.org/download_data.html).

Cliff AD, Ord JK (1981) *Spatial Processes: Models and Applications*. Pion Limited, London, UK.

Crutzen PJ, Andreae MO (1990) Biomass burning in the tropics: impact on atmospheric chemistry and biogeochemical cycles. *Science*, **250**, 1669–1678.

Dickson BG, Prather JW, Xu YG, Hampton HM, Aumack EN, Sisk TD (2006) Mapping the probability of large fire occurrence in northern Arizona, USA. *Landscape Ecology*, **21**, 747–761.

Dikshit O, Roy DP (1996) An empirical investigation of image resampling effects upon the spectral and textural information content of a high spatial resolution multispectral image. *Photogrammetric Engineering & Remote Sensing*, **62**, 1085–1092.

Edwards D (1984) Fire regimes in the biomes of South Africa. In: *Ecological Effects of Fire in South African Ecosystems*, Vol. 48 (eds Booysen P, Tainton NM), pp. 20–36. Springer-Verlag, Berlin, Germany.

FAO (2005) Global Livestock Distributions. <http://ergodd.200.ox.ac.uk/agaadat/index.htm> Vol. 2006.

Fritz H, Duncan P (1994) On the carrying capacity for large ungulates of African Savanna ecosystems. *Proceedings: Biological Sciences*, **256**, 77–82.

Frost PGH (1999) *Fire in southern African woodlands: origins, impacts, effects, and control*. Proceedings of an FAO meeting on public policies affecting forest fires. FAO Forestry Paper, 138.

GHRC (2003) Global Hydrology Resource Center (GHRC) at the Global Hydrology and Climate Center, LIS/OTD 0.5 Degree high resolution full climatology. <http://thunder.Asstc.nasa.gov/da> Huntsville, AL, USA.

Giglio L, Desclotres J, Justice CO, Kaufman YJ (2003) An enhanced contextual fire detection algorithm for MODIS. *Remote Sensing of Environment*, **87**, 273–282.

Gill AM (1975) Fire and the Australian flora: a review. *Australian Forestry*, **38**, 4–25.

Hansen MC, DeFries S, Townshend JRG, Carroll M, Cimiceli C, Sohlberg RA (2003) Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS vegetation continuous fields algorithm. *Earth Interactions*, **7**, 1–5.

Heyerdahl EK, Brubaker LB, Agee JK (2001) Spatial controls of historical fire regimes: a multiscale example from the interior west USA. *Ecology*, **82**, 660–678.

Higgins SI, Bond WJ, Trollope WSW (2000) Fire, resprouting and variability: a recipe for grass-tree coexistence in savanna. *Journal of Ecology*, **88**, 213–229.

Huffman GJ, Adler RF, Rudolph B, Schneider U, Keen P (1995) Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. *Journal of Climate*, **8**, 1284–1295.

IGBP (2000) Global Soil Data Products CD-ROM. International Geosphere Biosphere Program.

Justice C, Townshend J, Vermote E *et al.* (2002) An overview of MODIS land data processing and product status. *Remote Sensing of Environment*, **83**, 3–15.



- Kaufman YJ, Justice C, Flynn LP *et al.* (1998) Potential global fire monitoring from EOS-MODIS. *Journal of Geophysical Research*, **103**, 32215–32238.
- Keeley JE, Fotheringham CJ (2001) Historic fire regime in Southern California shrublands. *Conservation Biology*, **15**, 1536–1550.
- Keeley JE, Fotheringham CJ, Morais M (1999) Reexamining fire suppression impacts on Brushland fire regimes. *Science*, **284**, 1829–1833.
- Kirk M (1999) *Land Tenure, Technological Change and Resource Use: Transformation Processes in African Agrarian Systems*. Peter Lang Publishing, Frankfurt, Germany.
- Koenker R (2005) *Quantile Regression*. Cambridge University Press, Cambridge, UK.
- Laris PS (2005) Spatiotemporal problems with detecting and mapping mosaic fire regimes with coarse-resolution satellite data in savanna environments. *Remote Sensing of Environment*, **99**, 412–424.
- Legendre P, Legendre L (1998) *Numerical Ecology*. Elsevier Science, Amsterdam, the Netherlands.
- Lennon JJ (2000) Red-shifts and red herrings in geographical ecology. *Ecography*, **23**, 101–113.
- Luke RH, McArthur AG (1978) *Bushfires in Australia*. Australian Govt. Pub. Serv., Canberra, Australia.
- Lynch H, Renkin R, Crabtree R, Moorcroft P (2006) The influence of previous Mountain Pine beetle (*Dendroctonus ponderosae*) activity on the 1988 Yellowstone fires. *Ecosystems*, **9**, 1318–1327.
- Malaise FP (1978) The miombo ecosystem. In: *Tropical Forest Ecosystems: A State of Knowledge Report*. Vol. 14, pp. 589–606. UNESCO, Paris, France.
- Mayaux P, Bartholome E, Fritz S, Belward A (2004) A new land-cover map of Africa for the year 2000. *Journal of Biogeography*, **31**, 861–877.
- McArthur AG (1966) *Weather and grassland fire behaviour*. Forestry and Timber Bureau, (eds Department of National Development), Leaflet 100. pp. 23. Canberra, Australia.
- Mendelsohn JM (2002). *Atlas of Namibia: a portrait of the land and its people*. David Philip, Cape Town, South Africa.
- Miller J, Franklin J, Aspinall R (2007) Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling*, **202**, 225–242.
- Moritz MA (2003) Spatiotemporal analysis of controls on shrubland fire regimes: age dependency and fire hazard. *Ecology*, **84**, 351–361.
- Mulqueeny CM (2005) *Spatial and Temporal Variation of the Fire Regime in Mkuzi Game Reserve*. University of Witwatersrand, Johannesburg, South Africa.
- Openshaw S (1984) The modifiable areal unit problem. In: *Concepts and Techniques in Modern Geography*, Vol. 38. Geo Books, Norwich, UK.
- Patra PK, Ishizawa M, Maksyutov S, Nakazawa T, Inoue G (2005) Role of biomass burning and climate anomalies for land-atmosphere carbon fluxes based on inverse modeling of atmospheric CO<sub>2</sub>. *Global Biogeochemical Cycles*, **19**, GB3005, doi: 10.1029/2004GB002258.
- Prasad AM, Iverson LR, Liaw A (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, **9**, 181–199.
- Rejwan C, Collins NC, Brunner LJ, Shuter BJ, Ridgway MS (1999) Tree regression analysis on the nesting habitat of smallmouth bass. *Ecology*, **80**, 341–348.
- Roy D, Frost P, Justice C *et al.* (2005a) The Southern African Fire Network (SAFNet) regional burned area product validation protocol. *International Journal of Remote Sensing*, **26**, 4265–4292.
- Roy DP, Boschetti L, Justice C, Ju J (2008) The collection 5 MODIS burned area product – global evaluation by comparison with the MODIS active fire product. *Remote Sensing of the Environment*, **112**, 3690–3707.
- Roy DP, Jin Y, Lewis PE, Justice CO (2005b) Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sensing of Environment*, **97**, 137–162.
- Russell-Smith J, Yates C, Whitehead PJ *et al.* (2007) Bushfires 'down under': patterns and implications of contemporary Australian landscape burning. *International Journal of Wildland Fire*, **16**, 361–377.
- Sankaran M, Hanan NP, Scholes RJ *et al.* (2005) Determinants of woody cover in African savannas. *Nature*, **438**, 846–849.
- Saunders DA, Hobbs RJ, Margules CR (1991) Biological consequences of ecosystem fragmentation: a review. *Conservation Biology*, **5**, 18–32.
- Scholes RJ (2003) Convex relationships in ecosystems containing mixtures of trees and grass. *Environmental and Resource Economics*, **26**, 559–574.
- Scholes RJ, Archer SR (1997) Tree-grass interactions in savannas. *Annual Review of Ecological Systematics*, **28**, 517–544.
- Scholes RJ, Kendall J, Justice CO (1996) The quantity of biomass burned in southern Africa. *Journal of Geophysical Research*, **101**, 23667–23676.
- Sheuyange A, Oba G, Weladji RB (2005) Effects of anthropogenic fire history on savanna vegetation in northeastern Namibia. *Journal of Environmental Management*, **75**, 189–198.
- Silva JMN, Pereira JMC, Cabral AI, Sa ACL, Vasconcelos MJP, Mota B, Gregoire JM (2003) An estimate of the area burned in southern Africa during the 2000 dry season using SPOT-VEGETATION satellite data. *Journal of Geophysical Research – Atmospheres*, **108**, 8498.
- Spessa A, McBeth B, Prentice C (2005) Relationships among fire frequency, rainfall and vegetation patterns in the wet-dry tropics of northern Australia: an analysis based on NOAA-AVHRR data. *Global Ecology and Biogeography*, **14**, 439–454.
- Stott P (2000) Combustion in tropical biomass fires: a critical review. *Progress in Physical Geography*, **24**, 355–377.
- STRM (2005) Shuttle Radar Topography Mission 90 m digital topographic data Version 2 NASA.
- Thonicke K, Spessa A, Prentice C, Harrison SP, Carmona-Moreno C (submitted) The influence of vegetation, fire spread and fire behaviour on global biomass burning and trace gas emissions. *Global Change Biology*.
- Trollope WSW (1984) Fire in Savanna. In: *Ecological Effects of Fire in South African Ecosystems*, Vol. 48 (eds Booysen P, Tainton NM), pp. 40–52. Springer-Verlag, Berlin, Germany.

- Trollope WSW, Trollope LA, Hartnett DC (2002) Fire behaviour a key factor in the fire ecology of African grasslands and savannas. In: *Forest Fire Research & Wildland Fire Safety*. Millpress, Rotterdam, the Netherlands. pp. 1–15.
- UNEP (2002) *Africa Environment Outlook: Past Present and Future Perspectives*. EarthPrint, Hertfordshire, UK.
- UNEP-WCMC (2006) The World Database on Protected Areas UNEP.
- Unwin DJ (1996) GIS, spatial analysis and spatial statistics. *Progress in Human Geography*, **20**, 540–541.
- Van Wilgen BW, Govender N, Biggs HC, Ntsala D, Funda XN (2004) Response of Savanna fire regimes to changing fire-management policies in a large African National Park. *Conservation Biology*, **18**, 1533–1540.
- VanWilgen BW, Scholes RJ (1997) The vegetation and fire regimes of southern hemisphere Africa. In: *Fire in Southern African Savannas: Ecological and Atmospheric Perspectives* (eds Andreae MO, Goldammer JG, Lindsay K), pp. 27–46. Witwatersrand University Press, Johannesburg, South Africa.
- Venevsky S, Thonicke K, Sitch S, Cramer W (2002) Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study. *Global Change Biology*, **8**, 984–998.
- Verlinden A, Laamanen R (2006) Long term fire scar monitoring with remote sensing in northern Namibia: relations between fire frequency, rainfall, land cover, fire management and trees. *Environmental Monitoring and Assessment*, **112**, 231–253.
- Viegas DX (2004) Slope and wind effects on fire propagation. *International Journal of Wildland Fire*, **13**, 143–156.
- Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW (2006) Warming and earlier spring increases western U.S. forest wildfire activity. *Science*, **313**, 940–943.
- White F (1983) The vegetation of Africa, a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. *Natural Resource Research*, **20**, 1–356.
- Williams CA, Hanan N, Neff JC, Scholes RJ, Berry JA, Denning AS, Baker DF (2007) Africa and the global carbon cycle. *Carbon Balance and Management BioMed Central*, **2**, 1–13.
- Yates C, Russell-Smith J (2002) *An assessment of the accuracy of DOLA's Northern Australia NOAA-AVHRR Fire Affected Area (FAA) map products*. Australia State of the Environment 2nd technical paper series (Biodiversity), Department of the Environment and Heritage. Canberra, Australia.