CrossMark

RESEARCH ARTICLE

Water availability is a principal driver of large-scale land cover spatial heterogeneity in sub-Saharan savannahs

Christopher G. Marston (ORCID) · **David M. Wilkinson** · **Sally C. Reynolds** ·
Julien Louys · **Hannah J. O'Regan**

Abstract

Context The heterogeneous mosaic nature of African savannah vegetation is a key aspect of its ecology. This study evaluates mosaic distributions and characteristics across sub-Saharan Africa, investigating the environmental drivers of mosaic formation.

Objectives This study was designed to determine: (1) on a continental scale, how frequent are mosaics in savannahs? and (2) what are the key environmental drivers in the formation of mosaics?

Methods Landsat ETM+ satellite imagery was used to generate land-cover maps for 39 sample areas across sub-Saharan Africa. The spatial complexity of land-cover mosaics at 4628 savannah sub-sites was quantified, and modelled using random forests to identify the relative importance of environmental variables driving mosaic presence.

Results Only six sub-sites constituted a single land-cover class, illustrating that mosaic habitats are abundant at the scale analysed (19.6 km^2), although mosaic characteristics varied considerably. Results indicate precipitation is most important in influencing mosaic complexity, followed by evapotranspiration, temperature, lithology and distance to rivers. Fire and ecosystem engineer presence are of lesser importance at this study scale.

Conclusions Mosaics are ubiquitous in the African savannahs studied, their presence influenced by multiple environmental drivers, with water being key. The lower importance of fire and large mammal disturbance is likely resulting from these highly individualistic site-based process varying between sites, resulting in no single, coherent, across-Africa disturbance signal, and/or lack of detail in available data at

C. G. Marston (✉)
Department of Geography, Edge Hill University, Ormskirk, Lancashire L39 4QP, UK
e-mail: marstonc@edgehill.ac.uk

D. M. Wilkinson
School of Life Sciences, University of Lincoln, Joseph Banks Laboratories, Green Lane, Lincoln LN6 7DL, UK

D. M. Wilkinson · H. J. O'Regan
Department of Classics and Archaeology, School of Humanities, University of Nottingham, Nottingham NG7 2RD, UK

S. C. Reynolds
Faculty of Science and Technology, Bournemouth University, Poole BH12 5BB, UK

J. Louys
Australian Research Centre for Human Evolution, Environmental Futures Research Institute, Griffith University, Brisbane 4111, Australia

this scale. Therefore, large-scale determinants of savannah mosaics appear climate-driven. Under future global warming scenarios, African savannahs are likely to become more homogenous.

Introduction

Savannahs are one of the most important terrestrial biomes, covering around 20% of the global land surface (Shorrocks and Bates 2015). Although the exact definition of savannah is confusingly variable, the key aspect of almost all definitions is that it is a biome composed of a mix of trees and grass—usually C_4 grass—although tree-free grasslands are also occasionally defined as savannah (Archibold 1995; Torello-Raventos et al. 2013; Shorrocks and Bates 2015). Savannahs are widespread in sub-Saharan Africa (Shorrocks and Bates 2015), and, being 'home to one of the richest accumulations of mammals in the world' (Turner and Antón 2004, p. 37), are of considerable importance for nature conservation (Lawton 1998; Grünewald et al. 2016).

Characterised by the co-dominance of trees and grasses, savannah lies on a continuum between grassland and tropical forest (Torello-Raventos et al. 2013, Fig. 9). At a continental scale, climate is important in explaining the distribution of savannah. For example, analysis of 854 sites across Africa suggested that at lower rainfall levels savannah may be a stable system, but at higher rainfalls (> 784 mm mean annual precipitation) savannah may require periodic disturbance events to prevent succession to forest (Sankaran et al. 2005). Key disturbance mechanisms include fire and grazing by large herbivores such as elephants (*Loxodonta africana*) (Laws 1970; Dublin et al. 1990; Midgley et al. 2010; Daskin et al. 2016; Marston et al. 2017).

The savannah biome started to become widespread during the late Miocene (Kaya et al. 2018), probably in part driven by declining carbon dioxide levels, climate change and fire frequency (Beerling and Osborne 2006). Historically this expansion of savannah in Africa has been seen as playing an important role in human evolution—for example in the context of increased thermal stress on hominins (e.g. Newman 1970; Wheeler 1984; Ruxton and Wilkinson 2011). This emphasis on the selective importance of more open habitats is often referred to as the 'savannah hypothesis'—now viewed as of limited importance due to more recent evidence suggesting that adaptations such as bipedal stance evolved in more forested conditions (Domínguez-Rodrigo 2014). However, as Domínguez-Rodrigo (2014) points out, the original savannah hypothesis envisaged one end of a continuum of savannah vegetation, open grassland with few trees. In fact many African savannahs are patchy, or mosaic (here defined by two or more land cover classes, used hereafter as a descriptor of land cover heterogeneity), with very local variation in tree cover levels. Indeed, since the 1970s there has been considerable interest in the role of mosaics in human evolution literature (Reynolds et al. 2015). These mosaic environments are also of wider ecological interest in the context of ideas about source-sink populations, metapopulations, species richness, and macroecological processes (e.g. Hanski 1998; Nee 2007; Louys et al. 2011; Stein et al. 2014).

Given that the heterogeneous mosaic nature of African savannah vegetation is key for questions in contemporary ecology and conservation (e.g. Du Toit et al. 2014), understanding the origins of, and impacts on, the grazing niche (e.g. Louys and Faith 2015), and the drivers of human evolution (e.g. O'Regan et al. 2016), it is important to ask 'how frequent are mosaics in savannahs, and what factors drive their formation?' Ground-based monitoring is costly, time-consuming, limited in spatial and temporal coverage, and unfeasible on a continental-level scale. Instead, recent studies have investigated landscape patchiness, and its causes, using satellite remote sensing. For example, at an individual reserve level in South Africa, MacFadyen et al. (2016) looked at environmental heterogeneity (habitat mosaics) in the Kruger National Park and found that rainfall and seasonality were important drivers. This raises interesting questions about the role of past and future climate changes in the pattern of savannah mosaics.

Many studies investigate the drivers for mosaics at a reserve level (Scholtz et al. 2014; MacFadyen et al. 2016; Veldhuis et al. 2016), however spatial scale can be key in addressing ecological questions (May 1994). For example Gillson (2004) in a study of East African savannah suggested that the key processes

determining tree density varied with both spatial and temporal scales. This raises the possibility that the key drivers for mosaic patchiness in African savannahs at a continental scale may differ from those identified in more spatially restricted studies. Here we use a remote sensing approach to quantify the extent and nature of mosaics, and potential drivers of mosaic formation within savannahs across sub-Saharan Africa. Specifically we ask (1) on a continental scale, how frequent are mosaics in savannahs, and (2) what are the key environmental drivers in the formation and maintenance of mosaics?

Materials and methods

Study areas

To quantify mosaic heterogeneity in modern African savannahs, we analysed 39 Landsat ETM+ satellite image pairs across sub-Saharan Africa (Fig. 1). Image locations were determined using a random number generator to produce a series of x and y coordinates within sub-Saharan Africa using ArcMap 10.2. For each point, the Landsat ETM+ imagery archive of the United States Geological Survey (USGS) was searched to select suitable images—locations without suitable cloud-free imagery were discounted from further analysis. These discounted locations were mainly associated with rainforest, rather than savannah. This random sampling approach did result in non-savannah areas being present within the satellite image footprints; these non-savannah areas were excluded from analysis at a later stage (see below), to retain the savannah-only focus of this study. Image acquisition dates and path/row are displayed in Online Resource 1.

Methodology: satellite image analysis

African savannah vegetation often exhibits large contrasts between dry and wet seasons, with herbaceous vegetation generally only green during the rainy season with senescence occurring shortly afterwards, and most woody plants remaining photosynthetically active over larger parts of the year (Brandt et al. 2016). Single-date image analysis can have limitations in discriminating between woody and herbaceous vegetation that is spectrally similar at certain times of year (Marston et al. 2017). To overcome this, composites of

wet and dry season imagery were generated for each study location to improve vegetation discrimination based on their phenological differences, and additionally, identify land cover classes present only at certain times of the year such as seasonal water. Here, wet and dry season imagery for each site was determined visually by vegetation state (senescent in the dry season, flushed in the wet season), rather than by calendar date. Although the Landsat ETM+ 30 m spatial resolution precludes identification of individual trees and shrubs, Marston et al. (2017) illustrated that land cover classifications of African savannahs generated using Landsat ETM+ imagery are remarkably congruent with classifications of the same locations generated from very high resolution (VHR) WorldView-2 and IKONOS imagery (which can identify individual trees and shrubs), despite some spatial detail loss. Therefore, medium resolution Landsat ETM+ imagery is considered appropriate for broad-scale land cover mapping of heterogeneous African savannahs.

The satellite images underwent image geometric correction, cloud and cloud shadow masking, and atmospheric correction pre-processing steps to ensure data robustness before compositing the wet and dry season images into a single dual-date composite image (Morton et al. 2011). The dual-date composite images were then classified using a 75-class ISODATA unsupervised classification technique with each class assigned a land cover class label corresponding to the classification nomenclature in Table 1. This large number of classes was used to minimise the problem of split land cover class spectral clusters (Wayman et al. 2001), and has been demonstrated to be effective for mapping semi-natural environments (Marston et al. 2017). Unsupervised classification methods are well established for regional and global land cover mapping (Loveland et al. 2000), and were preferred here due to the highly heterogeneous nature of many of the study areas, and the scarcity of suitable spectrally pure training areas. The classification nomenclature was based on a modified version of the Global Land Cover 2000 Land Cover Map of Africa classification system (Mayaux et al. 2004). This classification also pays special attention to the forest—grassland gradient, and as with Torello-Raventos et al. (2013) it stratifies this gradient into five forest to grassland categories at 25% intervals of tree canopy cover (100–75%, 75–50%, 50–25%, 25–5% and 5–0%). The
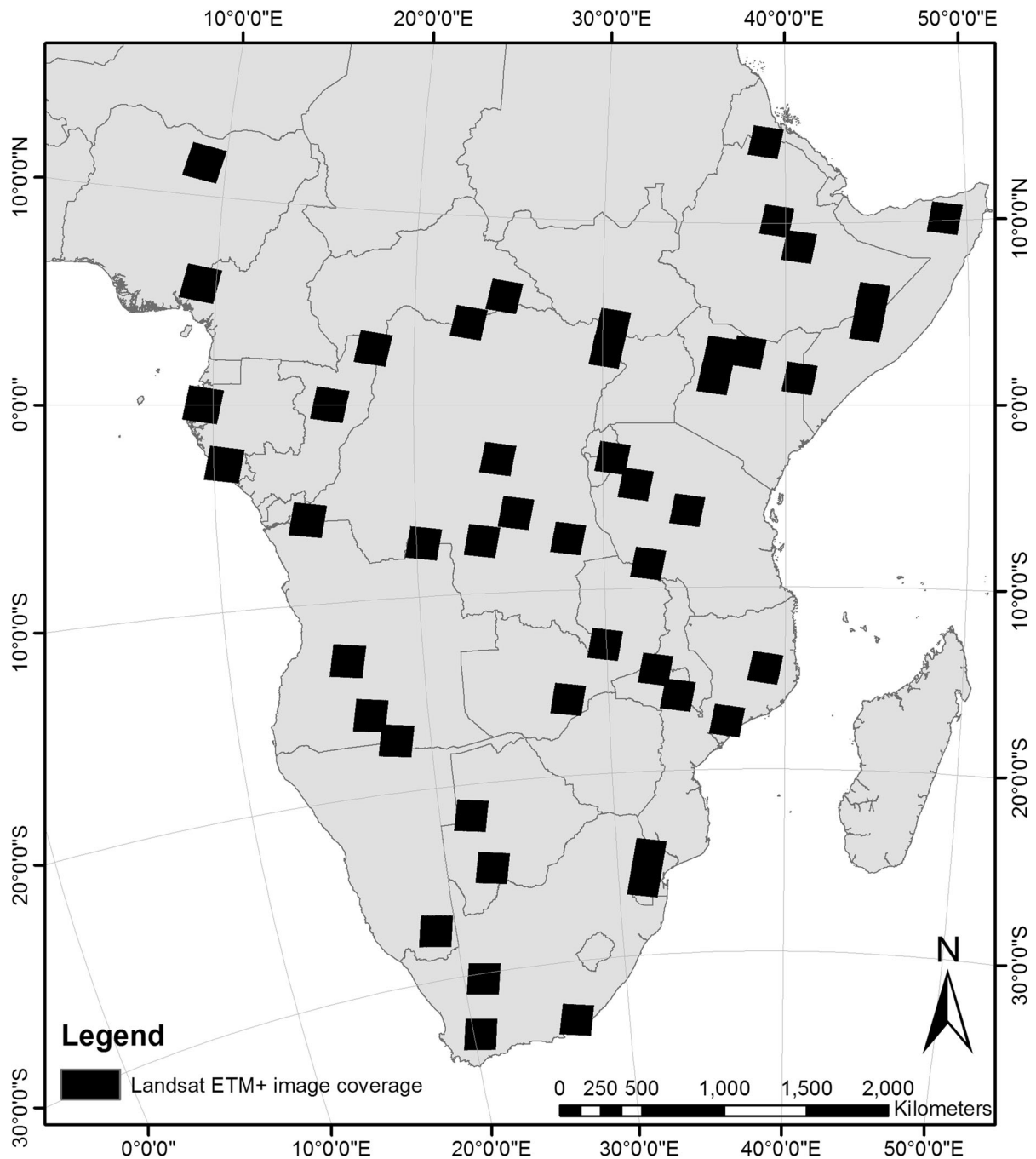
Fig. 1 Footprints of classified Landsat ETM+ images

latter two categories have here been amalgamated to form a 25–0% canopy grouping.

VHR satellite imagery of the study areas, available via public portals such as Google Earth, were used both as a validation dataset for image classification,

and for classification accuracy assessment. Using higher resolution imagery as a source of validation data for accuracy assessment of classifications derived from coarser resolution satellite imagery is an established technique (Duro et al. 2012). To ensure

Table 1 Land cover map classification nomenclature

General habitat	Land cover class
Woodland	Closed deciduous woodland (75–100% tree cover)
	Open deciduous woodland (50–75% tree cover)
Grassland	Continuous grassland (75–100% grassland)
	Discontinuous grassland (50–75% grassland)
Anthropogenic classes	Agriculture
	Built-up
	Closed coniferous woodland
Bare	Bare rock/soil
Freshwater	Permanent freshwater
	Seasonal freshwater
	Swamp
Coastal	Saltwater
	Mangrove
	Littoral sediment
	Supra-littoral sediment
	Saltmarsh
Semi-desert	Semi-desert
Ice and snow	Ice and snow
Sodic lake	Sodic lake

robustness, all validation points exhibiting suspected temporal change between the Landsat ETM+ and VHR reference data acquisition dates were disregarded. Additionally, field surveys conducted in the Kruger National Park, South Africa in July 2014 involved further ground truthing data collection (Marston et al. 2017). These field data were combined with the VHR-derived validation data, however given the logistical challenges of collecting ground-truthing data over such broad geographical areas, VHR-derived reference imagery provided the sole source of validation data for other sites. Accuracy assessment of the final classifications was performed, with accuracy figures shown in Online Resource 1.

Study area sub-sampling

Within the footprint of each classified image, a regular point grid with 10 km spacing between sample points was generated. Around each point, a 2.5 km radius circular buffer (corresponding to 19.6 km^2) was generated, with the area coverage of each land cover class present in the corresponding classification calculated using zonal statistics. This buffer size was chosen as relevant to investigating ideas about landscape use by early hominins (O'Regan et al. 2016),

making it relevant to many other medium to large sized mammals that are often a key focus of ecological and conservation related studies in African savannahs (Shorrocks and Bates 2015). Our previous study (O'Regan et al. 2016) used a subset of the current data to show that the number of habitat types were 'surprisingly scale invariant' (radius of buffers ranged from 694 to 13,000 m). The spatial configuration and distribution of land cover patches within these buffer areas was quantified using four landscape metric measures using FRAGSTATS (McGarigal et al. 2002). These are number of patches (NP; the total number of land cover patches within the buffer), patch richness (PR; the number of different land cover types i.e. 'habitats'), Shannons diversity index (SHDI; a metric that combines both the number and diversity of patches) and fractal dimension (FRAC; an index of complexity in mosaic pattern).

The extracted buffer data were quality checked to remove all buffers not fully contained within the image footprint, containing any cloud (due to uncertainty of the underlying land cover types), or > 80% water. As we are particularly interested in examining more 'natural' landscapes, buffers containing > 10% anthropogenic land cover classes (built-up, agriculture and closed coniferous woodland (coniferous

plantations)), were disregarded from further analyses. Finally, as this study focussed on savannah environments, the land cover classification for each buffer location as specified in the New Map of Standardized Terrestrial Ecosystems of Africa (Sayre et al. 2013) was used to subset the buffer area data to only savannah areas (corresponding to the 2. A Tropical Grassland, Savannah and Shrubland group). Of the original 14,340 buffers sites, 9802 were removed, leaving 4628 for further analysis.

Environmental drivers

To determine the causative drivers linked to mosaic habitat presence, the landscape metric (response) variables were modelled against a series of twelve potential drivers of savannah heterogeneity (Table 2) previously identified as influencing patchiness at different spatial and temporal scales. Partly abiotic drivers include rainfall (Sankaran et al. 2005; February et al. 2013), temperature and evapotranspiration (O'Brien et al. 1998, 2000), slope (O'Brien et al. 2000), lithology (Melzer et al. 2011), fire (Dublin et al. 1990; Midgley et al. 2010) and distance to rivers (O'Regan et al. 2016). Biotic drivers are centred around ecosystem engineers, and include elephants (Laws 1970; Dublin et al. 1990; Guldemond and van Aarde 2008), white rhinoceros (*Ceratotherium simum*) (Waldram et al. 2008; Cromsig and te Beest 2014), hippopotamuses (*Hippopotamus amphibius*) (Lock 1972; McCarthy et al. 1998), porcupines (*Hystrix* spp.) (Yeaton 1988), and mole rats (Honeycutt 2016). The potential drivers were modelled using random forests against the derived PR, NP, SHDI and FRAC landscape metric variables to identify their relative importance, following methods previously applied by Veldhuis et al. (2016), and performed in R using the 'randomForests' package (Liaw and Wiener 2002). Random forests are a machine learning approach to classification, very suitable for complex, non-linear ecological datasets (Cutler et al. 2007). Their use in ecology is becoming more widespread and recently they have been applied to a range of ecological datasets such as African savannah vegetation at game park scale (Veldhuis et al. 2016), landscape dynamic influences on disease vectors (Marston et al. 2014, 2016), and trait analysis in plants (Bergmann et al. 2017). The statistical significance of the relative importance values of the predictor variables are also evaluated using a permutation-based random forest approach. This generates a large number of random forest models to obtain the probability distributions of the relative importance measures of the predictors, then quantifies how rarely the original relative importance measure of each predictor is obtained by chance. In the form we have used here (i.e. 'statistically reinforced') random forests are considered to have important advantages over more classical statistical approaches for detecting non-linear relationships and higher-order interactions in complex datasets (Ryo and Rillig 2017). Further information on random forests is contained in Online Resource 2.

Results

How frequent are mosaics in savannahs?

To examine the presence and nature of mosaics (here defined as two or more land cover classes in a buffer), the proportional coverage of the land cover classes for the 4628 retained buffers is examined. The commonest land cover classes were open deciduous woodland (26.5% of the land cover area), closed deciduous woodland (24.5%), discontinuous grassland (18.4%), semi-desert (15.9%), continuous grassland (8.5%), and bare (4.0%). Swamp (0.3%), permanent freshwater (0.2%), and seasonal freshwater (0.1%) were also present, plus anthropogenic land cover classes of agriculture (1.2%) and built-up (0.1%) (Online Resource 3).

The variability in PR, NP, SHDI and FRAC metrics values is illustrated in Online Resource 4. Crucially, only six of 4628 buffers have a single land cover class present (open deciduous woodland = 1, closed deciduous woodland = 2, semi-desert = 3), illustrating that mosaic habitats are ubiquitous in African savannahs at the scale used in this analysis. However, the nature of these mosaic habitats varies considerably in terms of number of land cover classes present, and the proportional coverage make-up of these land cover types. In particular, there is huge variability in NP (min = 1, max = 3611, mean = 931.3, median = 854.5, variance = 302108.6). This variance is significantly different to PR, SHDI and FRAC when assessed using Brown-Forsyth tests (Online Resource 5), although once outliers are removed, significant differences exist only between NP and PR and FRAC. All other metrics

Table 2 Continental-level variables and data sources

Group	Dataset	Source
Topography	Slope—standard deviation of slope values across a 2.5 km radius buffered area.	Derived from ASTER Global Digital Elevation Map
Ecosystem engineer presence/absence	Naked mole rat (*Heterocephalus glaber*)	International Union for Conservation of Nature (IUCN) Red List of Threatened Species (IUCN 2018)
	Cape mole rat (*Georychus capensis*)	
	Silvery mole rat (*Heliophobius* spp.)	
	Common mole rat (*Cryptomys* spp.)	
	Cape dune mole rat (*Bathyergus* spp.)	
	White rhinoceros (*Ceratotherium simum*)	
	Hippopotamus (*Hippopotamus amphibius*)	
	Elephant (*Loxodonta africana*)	
	Cape Porcupine (*Hystrix africaeaustralis*)	
	Crested Porcupine (*Hystrix cristata*)	
Lithology	Lithology	A new map of standardized terrestrial ecosystems of Africa
Climate	Mean total annual precipitation (1950–2000)	WorldClim—global climate data
	Mean annual temperature (1950–2000)	
Evapotranspiration	Global potential evapotranspiration	CGIAR Consortium for Spatial Information (CGIAR-CSI) Global Aridity Index
Mean fires per year	ATSR World Fire Atlas	European Space Agency ATSR World Fire Atlas
Distance to rivers	Distance to nearest river	National Oceanic and Atmospheric Administration (NOAA)

For modelling purposes, mole rat spp. and porcupine spp. are amalgamated into single mole rat and porcupine presence/absence variables respectively

show statistically indistinguishable levels of variance with outliers present or removed.

SHDI frequency distribution violin plot (Online Resource 4) illustrates the distribution of diversity values across the buffers (min = 0.00, max = 2.00, mean = 0.91, median = 0.95, variance = 0.15), and demonstrate how both the richness and evenness of land cover patches varies. These results are consistent with the PR results (min = 1, max = 12, mean = 5.56, median = 6, variance = 3.94) in showing considerable variability in patterns of landscape heterogeneity throughout the dataset. Fractal dimension values showed lower levels of variance (min = 1.01, max = 1.42, mean = 1.26, median = 1.27, variance = 0.003), although these were only significantly different to NP (Table S2). This suggests that the distribution of mosaics within the study areas are consistent regardless of the number of patches in the landscape.

What are the key environmental drivers of mosaics?

Random forest analysis was used to establish the relative importance of the environmental drivers in relation to the metric values, and explained 74.58% (PR), 67.4% (NP), 63.7% (SHDI) and 49.42% (FRAC) of variance respectively. Table 3 displays the percentage increase in mean square error when values for the chosen variable is randomly assigned throughout the dataset, scaled to assign the most important predictor a value of 100 (Veldhuis et al. 2016). High mean square error shows an increased importance for that variable. This analysis indicates that, overall, total annual precipitation is the most important variable influencing landscape metric values, with evapotranspiration second, mean annual temperature third, followed by lithology and distance to rivers (Table 3). Slope, mean fires per year and ecosystem engineers are of lesser importance. The rank order of importance of

Table 3 Relative importance values of the predictor values for explaining NP, PR, SHDI and FRAC

Variable	NP	PR	SHDI	FRAC
Slope	66.6 (6)	74.9 (6)	63.4 (6)	51.7 (6)
Mole rat spp.	58.3 (7)	49.8 (7)	40.4 (8)	35.3 (8)
White Rhino	42.7 (9)	36.5 (10)	39.8 (9)	33.2 (9)
Hippopotamus	36.7 (11)	40.1 (8)	35.4 (11)	21.4 (12)
Elephant	39.9 (10)	25.1 (12)	39.7 (10)	31.9 (10)
Porcupine spp.	28.3 (12)	33.2 (11)	24.8 (12)	22.5 (11)
Lithology	79.2 (4)	82.0 (4)	74.4 (4)	62.7 (4)
Total annual precipitation	100.0 (1)	98.3 (2)	100.0 (1)	100.0 (1)
Mean annual temperature	83.8 (3)	79.8 (5)	74.9 (3)	66.9 (3)
Evapotranspiration	85.5 (2)	100.0 (1)	78.7 (2)	76.5 (2)
Mean fires per year	43.8 (8)	38.2 (9)	52.2 (7)	35.4 (7)
Distance to rivers	77.0 (5)	88.3 (3)	68.1 (5)	61.6 (5)

Relative importance values were obtained by assigning the most important predictor a value of 100, and scaling the others appropriately (Veldhuis et al. 2016). Variable importance rankings are displayed for each variable in parentheses

all driver variables for each of the NP, PR, SHDI and FRAC metrics were shown to be statistically significant at the 95% confidence level by the permutation-based random forest models ($P < 0.05$) (see Online Resource 2). Correlations between mean fires per year, precipitation and mean annual temperature were also calculated with scatterplots presented in Online Resource 6. Pearson's correlations and p-values (fires—precipitation: correlation = 0.24, p-value = < 0.001; fires—temperature: correlation = 0.05, p-value = < 0.001) indicated that although there were statistically significant relationships between fire and precipitation and temperature, correlations were relatively weak.

Random forest partial dependence plots illustrate the nature of the effects of drivers on the metric values (Figs. 2, 3). Total annual precipitation consistently exhibits a hump-shaped curve for all four metrics, with peaks in metric values between precipitation levels of approximately 350–1600 mm. Evapotranspiration illustrated a generally linear pattern for NP, PR, and SHDI, with consistently high metrics values recorded at evapotranspiration levels of approximately 1550 mm/month or below, followed by a rapid decline between 1550 and 2000, then consistently low levels above 2000. FRAC values were consistently at moderate levels below 1700 mm/month, then increased to a peak at approximately 1850, before rapidly declining to low levels above 2100 mm/month. This peak corresponds with the smaller peak in SHDI, and the trough in PR, indicating that there are not many different land cover types present at this point in the distribution, but they are

covering the landscape quite evenly. Mean annual temperature also displays a generally consistent hump-shaped distribution for all four metrics, with low values below 18 °C increasing and peaking at around 23–27 °C. Above 27 °C, FRAC remains consistently high, PR shows a slight reduction, and NP and SHDI exhibit a larger decline. The influence of distance to rivers varied, with peaks of PR and SHDI values close to rivers before a rapid drop-off at approximately 10 km from a river, then a more gradual decline before levelling off at around 100 km. NP showed a similar initial peak then drop-off in values before climbing to a second peak at approximately 60 km, before reducing slightly and levelling off. FRAC exhibited a converse pattern with initial low values in close proximity to the river, rising rapidly to high values between 10 and 90 km, followed by a further rapid reduction in values > 90 km from a river.

Standard deviation (SD) of slope values across the buffer areas exhibited differing patterns for the four landscape metrics. NP and FRAC both showed initial peaks where SD slope = 0, then showed rapid reductions in metric values as SD slope increased, with FRAC remaining low while NP gradually increased once again. PR initially was very low before exhibiting an initially rapid, then more gradual increase as SD slope increased. SHDI showed considerable variability, with an initial peak then reduction in metric values, before fluctuating at lower levels. Mean fires per year showed a consistent pattern for NP, PR and SHDI, with very low values at 0 fires per year before a rapid increase between 0 and 0.2, before another
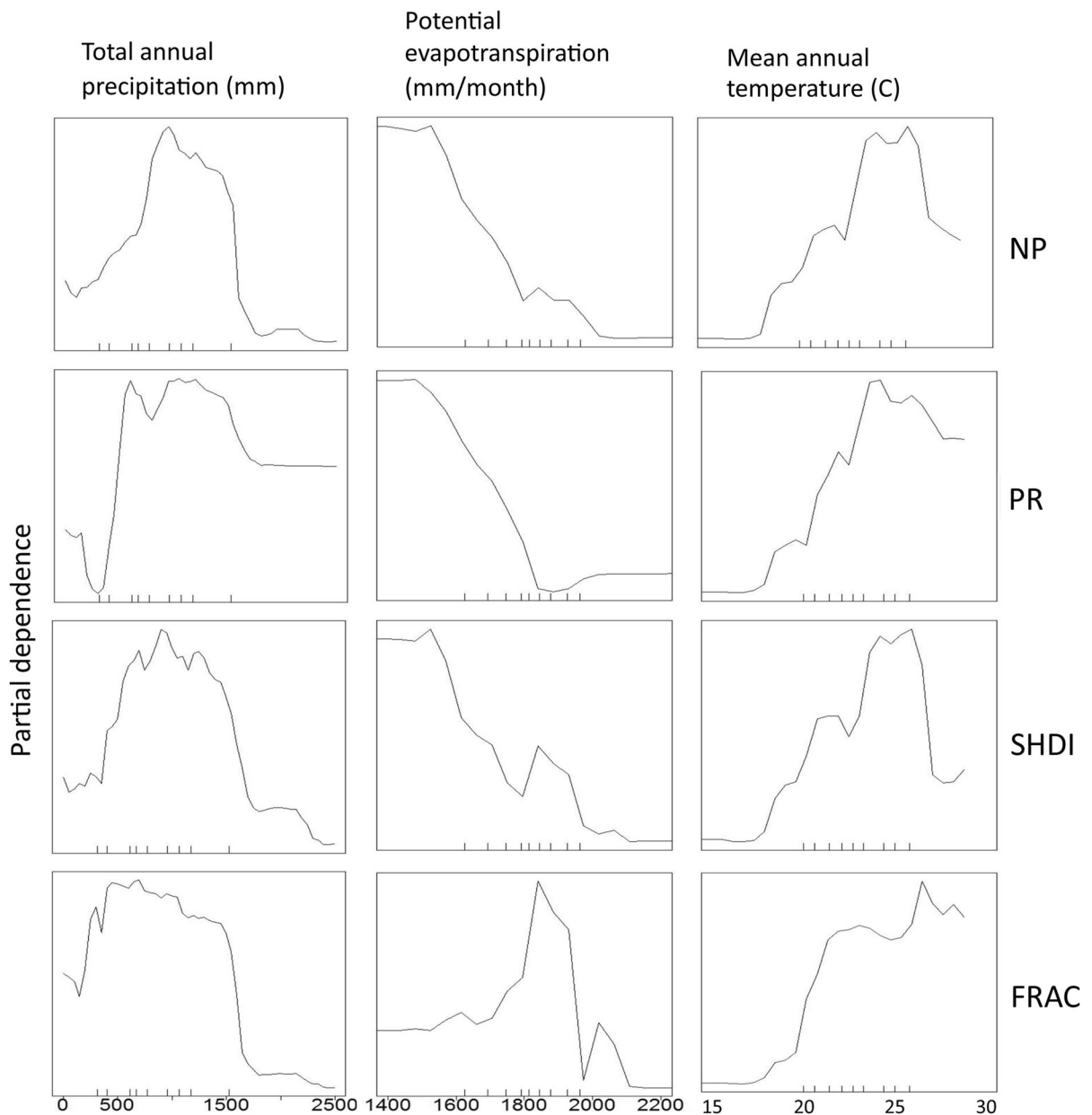
Fig. 2 Partial dependence plots for the PR, NP, SHDI and FRAC landscape metric (response) variables, and the total annual precipitation, potential evapotranspiration and mean annual temperature variables

gradual reduction for PR and SHDI between approximately 0.5 and 0.9, with metric values remaining steady above 1.0 mean fires per year. NP values remained high above 0.2 mean fires per year. FRAC differed, with initially high values at 0 fires per year, before a steady reduction to very low levels at 1.0

mean fires per year, with values remaining very low at higher levels of fire incidence. Partial dependence plots for ecosystem engineer species and lithology variables displayed little variability between species presence or absence, or between lithological classes (see Online Resources 7 and 8).
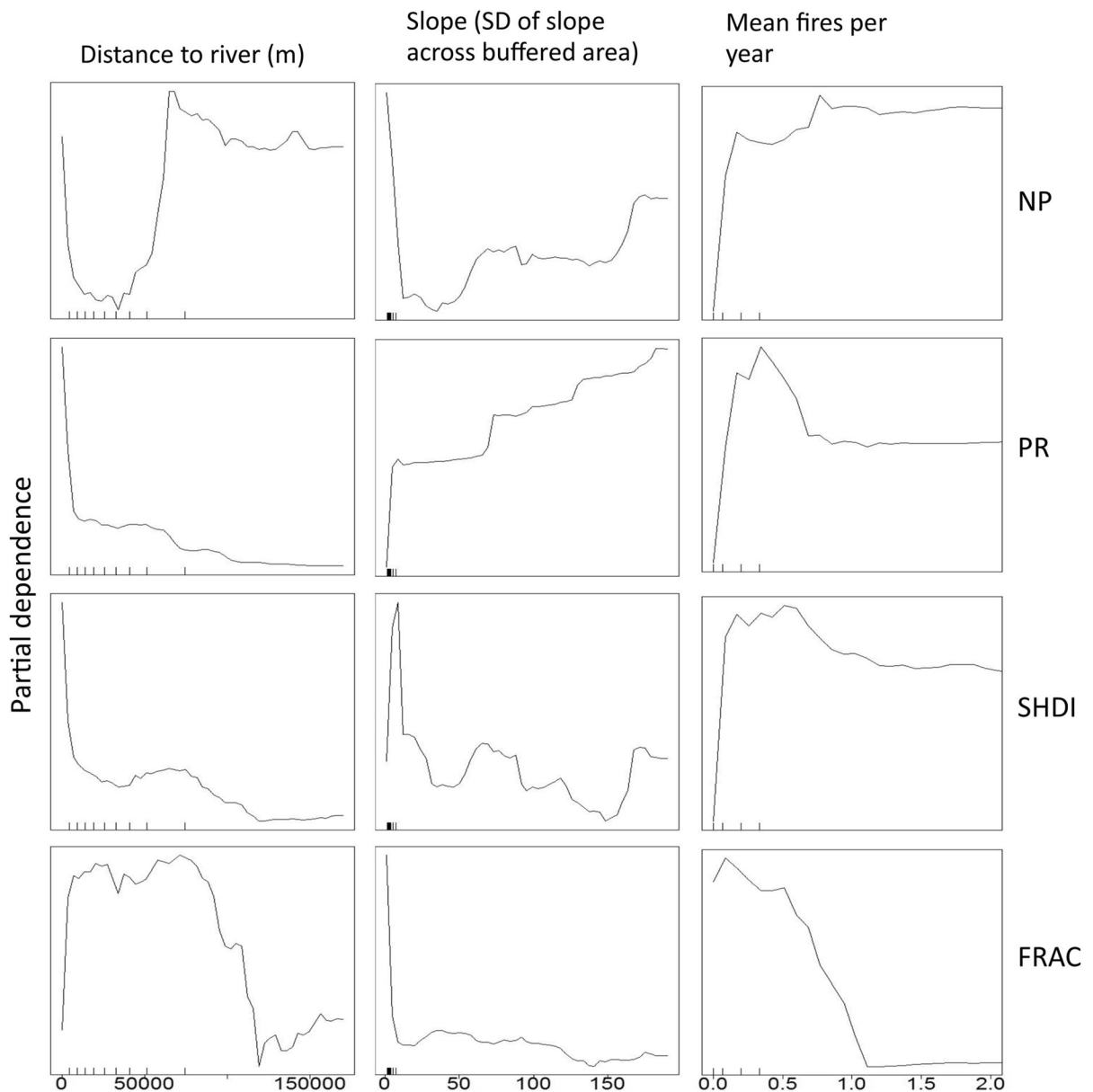
Fig. 3 Partial dependence plots for the PR, NP, SHDI and FRAC landscape metric (response) variables, and the distance to river, slope and mean fires per year variables

Discussion

Our first question, 'how frequent are mosaics in savannahs?' is easily answered by our analysis. Only 0.1% of our buffers were composed of a single land cover type. Mosaics are effectively ubiquitous in the African savannahs that we sampled at a scale of analysis suitable for medium to large mammals (a buffer of 19.6 km^2). As described in the Introduction

savannahs are normally defined as a mix of grass and woody vegetation—this could be thought to imply that a mosaic nature is a forgone conclusion. However, it might be the case that at the scale of analysis in our study some buffers could have been composed of just one habitat type (both of our woodland types, along with our two grassland types, can be formed from a mix of woody vegetation and grass in varying proportions. See Table 1). Our results clearly show

that in modern African savannahs it is an extremely rare for only one habitat type to exist at our scale of analysis. Within mosaic savannah environments, the number of patches varied the most across our metrics, although the number of different types of patches and their pattern of distribution across the landscape also showed considerable variance.

The answer to our second question 'What are the key environmental drivers in the formation and maintenance of mosaics?' is more nuanced. Clearly the formation of mosaics is a process that happens over time, and our study is a snapshot of the current pattern and associated potential 'drivers'. In the absence of appropriate historical time-series data we are effectively using a space for time substitution approach— common in ecology where, as here, other options are restricted, but not without a series of well-known issues and caveats (Pickett 1989). Across the African regions sampled in this study, total annual precipitation and evapotranspiration are the most important predictors of the extent of mosaics in savannahs. Total annual precipitation shows a hump-shaped pattern on the partial dependence plots, which seems to represent the rainfall levels associated with unstable savannah as described by Sankaran et al. (2005). This appears to be ultimately about the presence of trees—an important part of savannah mosaics. Distance to rivers has a similar mechanism—the presence of water leading to an increased likelihood of trees by rivers. The most successful predictors are all aspects of climate, followed by geomorphological factors (especially lithology) that likely also influence the local micro-climate and/or water content of soils. Similar relationships with lithology are well known from site specific studies—as in the Kruger National Park, South Africa (Scholtz et al. 2014). Therefore, all the key predictors in the random forest analysis are ones that facilitate tree growth, the presence of patches of trees being a crucial aspect of African savannah mosaics—unless the trees develop to the point that the mosaic is replaced by extensive wooded habitat. This has implications for making semi-quantitative estimates of mosaic extents in the past (for example in the content of human evolution), as climate and lithology are likely easier to estimate than other potential factors such as disturbance by large herbivores.

A surprising aspect of our results is the relative lack of importance given to disturbance by fire and large mammal grazers and browsers—although p-values need interpreting with caution, the permutation-based significance of our random forest analyses provides some formal statistical support that this is a real pattern in need of explanation. The literature on African savannah vegetation provides extensive evidence for the importance of such disturbance in maintaining a mosaic vegetation that does not succeed to forest (e.g. Laws 1970; Dublin et al. 1990; Midgley et al. 2010; Pringle et al. 2014; Shorrocks and Bates 2015). There are two likely explanations for the lower importance of disturbance in our analysis. One possibility is that the data sets used in our analysis of herbivores effects are not detailed enough to pick up these effects as they focus on individual ecosystem engineers rather than attempts to quantify overall grazing or browsing levels. Alternatively it seems possible that at the scale of a site/individual reserve, the disturbance effects of large mammals (especially elephants) and fire are important for mosaic presence, but at the scale of sub-Saharan Africa they are of lesser importance as key disturbance mechanisms will vary from site to site. Therefore, there is no single, coherent, across-Africa disturbance signal, and the large-scale determinant of savannah mosaics appears to be climate-driven.

Support for this hypothesis comes from a study of savannah at Hluhluwe-iMfolozi Park, South Africa, which using similar methods, found both rainfall and fire frequency to be key drivers of savannah mosaics (Veldhuis et al. 2016). Large herbivores effects were of lesser magnitude, and more complex than fire or rainfall. Recently in an analysis of woody plant encroachment across sub-Saharan Africa Venter et al. (2018) suggested that grazing mammals can have 'contradictory' effects—with high numbers sometimes increasing, and sometimes decreasing woody plant encroachment. Elephants can also have a more complex effect on woody vegetation than is suggested by the common assumption that they always lead to a decrease in woody vegetation (Kohi et al. 2011). In addition sparse data availability precluded the modelling of some ecosystem engineer species including termites, which have been identified as influential in modifying vegetation patterns (Bonachela et al. 2015). Limitations in the ATSR World Fire Atlas (WFA) data used to calculate mean fires per year in this study are also acknowledged. The WFA data is generated from 1 km resolution data from the ERS-2 satellite, offering a revisiting period of three days at

the equator. Although capable of detecting small burning areas of 0.01 ha at 800 °K (Mota et al. 2006), non-detection of fires occurring between satellite revisit times is acknowledged to underestimate fire activity (Kasischke et al. 2003), however this limitation would be consistent across all study areas, so respective fire occurrence values for different sites are therefore directly comparable. The calculation of mean number of fires per year over the duration of the WFA data set (November 1995–June 2004), rather than assessment of individual annual totals, also mitigates against the effect of individual-year anomalous fire occurrences from factors such as the El Niño Southern Oscillation (ENSO). Therefore the WFA data is the most appropriate dataset for monitoring fire occurrence for this time period over such a broad area.

At the across-Africa scale of this study, water is a key determinant of the mosaic characteristics of savannahs. This raises important implications for the future of savannah mosaics and their conservation importance in the context of climate change, with our random forest results allowing speculation on indicative future patterns. The predicted precipitation reductions in southern Africa and increases in eastern Africa presented by Shongwe et al. (2009, 2011) is likely to result in increased homogeneity in southern African savannahs, with increased grassland dominance (Sankaran et al. 2005). In eastern African savannahs, heterogeneity is likely to be maintained at 600–1400 mm precipitation levels, but will decrease at > 1400 mm, with increasing tree cover at these higher precipitation levels (Sankaran et al. 2005). Higher mean annual temperatures would also see increased savannah heterogeneity between 18 and 23 °C, while at extremes of > 25 °C the pattern will reverse, resulting in reduced heterogeneity. Consequently, for savannahs at the upper extremes of the mean annual temperature range, under global warming, our results suggest reductions in the number and types of patches in savannahs, and the surviving patches will be highly clumped. Other variables potentially influenced by future climate change include increased fires per year (as a consequence of increasing temperature) which is unlikely to affect NP, and show minor reductions in SHDI and PR, but will dramatically affect FRAC, where more than one fire a year will dramatically reduce the fractal dimension of patch distribution, essentially homogenising the landscape. Reduced rainfall causing proportionally greater

reductions in surface drainage (De Wit and Stankiewicz 2006), will also result in reductions in heterogeneity in close proximity to rivers affected by reduced flows. Increasing CO_2 over time (not included as a variable in our modelling as we investigated spatial rather than temporal variability) has also been identified as a driver of woody encroachment reducing heterogeneity in African savannahs (Bond and Midgley 2012; Marston et al. 2017).

Here, we have shown that satellite remote sensing can be applied successfully to monitor and quantify the mosaic nature of African savannahs. This has enabled, for the first time, broad-scale investigation of mosaics in savannahs at study areas across sub-Saharan Africa. Importantly, we find that mosaics are ubiquitous in the areas studied, and that the formation and maintenance of these mosaics is influenced by a number of environmental drivers, with water ultimately being the key driver. The role of disturbance by mega-herbivores does not emerge as a strong driver of mosaics at the scale used in our analysis, suggesting disturbance is a highly individualistic process for each region/site. Our findings also suggest that significant changes to abiotic drivers examined here, under scenarios of future global warming, will significantly impact on the mosaic nature of savannahs. Importantly, our study indicates that African savannahs, a key terrestrial biome of significant ecological importance, will become more homogenous under the increased temperatures, increased and decreased precipitation, and increased fires predicted to accompany future climate change. In addition these patterns also allow suggestions to be made about past mosaics in the context of Quaternary climate changes.

Author contributions HJO'R, DMW, SCR and JL conceived the ideas; CGM, HJO'R and DMW collected the field data;

CGM, SCR, and JL analysed the remote sensing data. CGM and DMW led the writing, with contributions from all authors. All authors give final approval for publication.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Archibold OW (1995) Ecology of world vegetation. Chapman and Hall, London

Beerling DJ, Osborne CP (2006) The origin of the savanna biome. Glob Change Biol 12:2023–2031

Bergmann J, Ryo M, Prati D, Hempel S, Rillig M (2017) Root traits are more than analogues of leaf traits: the case for diaspore mass. New Phytol 216:1130–1139

Bonachela JA, Pringle RM, Sheffer E, Coverdale TC, Guyton JA, Caylor KK, Levin SA, Tarnita CE (2015) Termite mounds can increase the robustness of dryland ecosystems to climatic change. Science 347:651–655

Bond WJ, Midgley GF (2012) Carbon dioxide and the uneasy interactions of trees and savannah grasses. Philos T R Soc B 367:601–612

Brandt M, Hiernaux P, Tagesson T, Verger A, Rasmussen K, Diouf AA, Mbow C, Mougin E, Fensholt R (2016) Woody plant cover estimation in drylands from Earth Observation based seasonal metrics. Remote Sens Environ 172:28–38

Cromsigt JPGM, te Beest M (2014) Restoration of a megaherbivore: landscape-level impacts of white rhinoceros in Kruger National Park, South Africa. J Ecol 102:566–575

Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. Ecology 88:2783–2792

Daskin JH, Stalmans M, Pringle RM (2016) Ecological legacies of civil war: 35-year increase in savannah tree cover following wholesale large-mammal declines. J Ecol 104:79–89

De Wit M, Stankiewicz J (2006) Changes in surface water supply across Africa with predicted climate change. Science 311:1917–1921

Domínguez-Rodrigo M (2014) Is the "Savanna Hypothesis" a dead concept for explaining the emergence of the earliest hominins? Curr Anthropol 55:59–81

Du Toit JT, Skarpe C, Moe SR (2014) Elephants and heterogeneity in Savanna landscapes. In: Skarpe C, Du Toit JT, Moe SR (eds) Elephants and savanna woodland ecosystems: a study from Chobe National Park, Botswana. John Wiley, London, pp 289–298

Dublin HT, Sinclair ARE, McGlade J (1990) Elephants and fire as causes of multiple stable states in the Serengeti-Mara woodlands. J Anim Ecol 59:1147–1164

Duro DC, Franklin SE, Dube MG (2012) Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. Int J Remote Sens 33:4502–4526

February EC, Higgins SI, Bond WJ, Swemmer L (2013) Influence of competition and rainfall manipulation on the growth response of savanna trees and grasses. Ecology 94:1155–1164

Gillson L (2004) Evidence of hierarchical patch dynamics in an East African savanna? Landscape Ecol 19:883–894

Grünewald C, Schleuning M, Böhning-Gaese K (2016) Biodiversity, scenery and infrastructure: factors driving wildlife tourism in an African savannah national park. Biol Conserv 201:60–68

Guldemond R, van Aarde R (2008) A meta-analysis of the impact of African elephants on savanna vegetation. J Wildl Manage 72:892–899

Hanski I (1998) Metapopulation dynamics. Nature 396:41–49

Honeycutt RL (2016) Family Bathyergidae (African mole rats). In: Wilson DE, Lacher TE Jr, Mittermeier RA (eds) Handbook of the mammals of the world, vol 6. Lynx Edicions, Barcelona, pp 352–370

IUCN (2018) The IUCN Red List of Threatened Species. Version 2018-1. http://www.iucnredlist.org

Kasischke ES, Hewson JH, Stocks B, van der Werf G, Randerson J (2003) 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

Kaya F, Bibi F, Zliobite I, Eronen JT, Hui T, Fortelius M (2018) The rise and fall of the Old World savannah fauna and the origins of the African savannah biome. Nat Ecol Evol 2:241–246

Kohi EM, de Boer WF, Peel KJS, Slotow R, van der Waal C, Heitkönig IMA, Skidmore A, Prins HTH (2011) African elephants *Loxodonta africana* amplify browse heterogeneity in African savanna. Biotropica 43:711–721

Laws RM (1970) Elephants as agents of habitat and landscape change in East Africa. Oikos 21:1–15

Lawton JH (1998) Green tourism and nature's services. Oikos 82:3–4

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2:18–22

Lock JM (1972) The effects of hippopotamus grazing on grasslands. J Ecol 60:445–467

Louys J, Faith JT (2015) Phylogenetic topology mapped onto dietary ecospace reveals multiple pathways in the evolution of the herbivorous niche in African Bovidae. J Zool Syst Evol Res 53:140–154

Louys J, Meloro C, Elton S, Ditchfield P, Bishop LC (2011) Mammal community structure correlates with arboreal heterogeneity in faunally and geographically diverse

habitats: implications for community convergence. Glob Ecol Biogeogr 20:717–729

Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu Z, Yang L, Merchant JW (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int J Remote Sens 21:1303–1330

MacFadyen S, Hui C, Verburg PH, Teeffelen AJA (2016) Quantifying spatiotemporal drivers of environmental heterogeneity in Kruger National Park, South Africa. Landscape Ecol 31:2013–2029

Marston CG, Danson FM, Armitage RP, Giraudoux P, Pleydell DRJ, Wang Q, Qiu J, Craig PS (2014) A random forest approach to describing Echinococcus multilocularis reservoir Ochotona spp. presence in relation to landscape characteristics in western China. Appl Geogr 55:176–183

Marston CG, Giraudoux P, Armitage RP, Danson FM, Reynolds S, Wang Q, Qiu J, Craig PS (2016) Vegetation phenology and habitat discrimination: impacts for E. multilocularis transmission host modelling. Remote Sens Environ 176:320–327

Marston CG, Aplin P, Wilkinson DM, Field R, O'Regan HJ (2017) Scrubbing up: multi-scale investigations of woody encroachment in a Southern African savannah. Remote Sens 9:419

May RM (1994) The effects of spatial scale on ecological questions and answers. In: Edwards PJ, May RM, Webb NR (eds) Large-scale ecology and conservation biology. Blackwell, Oxford, pp 1–17

Mayaux P, Bartholom E, Fritz S, Belward A (2004) A new land-cover map of Africa for the Year 2000. J Biogeogr 31:861–877

McCarthy TS, Ellery N, Bloem A (1998) Some observations on the geomorphological impact of hippopotamus (Hippopotamus amphibious L.) in the Okavango Delta, Botswana. Afr J Ecol 36:44–56

McGarigal K, Cushman SA, Neel MC, Ene E (2002). FRAGSTATS: Spatial pattern analysis program for categorical maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. www.umass.edu/landeco/research/fragstats/fragstats.html

Melzer SE, Chadwick OA, Hartshorn AS, Khomo LM, Knapp AK, Kelly EG (2011) Lithologic controls on biogenic silica cycling in South African savanna ecosystems. Biogeochemistry 108:317–334

Midgley JJ, Lawes MJ, Chamaillé-Jammes S (2010) Savanna woody plant dynamics: the role of fire and herbivory, separately and synergistically. Aust J Bot 58:1–11

Morton D, Rowland C, Wood C, Meek L, Marston C, Smith G, Wadsworth R, Simpson IC (2011) Final report for LCM2007—the new UK land cover map. Countryside survey technical report No 11/07. NERC/Centre for. Ecol Hydrol 112:pp

Mota B, Pereira JMC, Oom D, Vasconcelos M, Schultz M (2006) Screening the ESA ATSR-2 world fire atlas (1997–2002). Atmos Chem Phys 6:1409–1424

Nee S (2007) Metapopulations and their spatial dynamics. In: May RM, McLean AR (eds) Theoretical ecology, 3rd edn. Oxford University Press, Oxford, pp 35–45

Newman RW (1970) Why man is such a sweaty and thirsty naked animal: a speculative review. Hum Biol 42:12–27

O'Brien EM, Whittaker RJ, Field R (1998) Climate and woody plant diversity in southern Africa: relationships at species, genus and family levels. Ecography 21:495–509

O'Brien EM, Field R, Whittaker RJ (2000) Climatic gradients in woody plant (tree and shrub) diversity: water-energy dynamics, residual variation, and topography. Oikos 89:588–600

O'Regan HJ, Wilkinson DM, Marston CG (2016) Hominin home ranges and habitat variability: exploring modern African analogues using remote sensing. J Archaeol Sci Rep 9:238–248

Pickett STA (1989) Space-for-Time Substitution as an Alternative to Long-Term Studies. In: Likens GE (ed) Long-term studies in ecology. Springer, New York, pp 110–135

Pringle RM, Goheen JR, Palmer TM, Charles GK, DeFranco E, Hohbein R, Ford AT, Tarnita CE (2014) Low functional redundancy among mammalian browsers in regulating an encroaching shrub (Solanum campylacanthum) in African savanna. Proc R Soc B 281:20140390

Reynolds SC, Wilkinson DM, Marston CG, O'Regan HJ (2015) The 'mosaic habitat' concept in human evolution: past and present. Trans R Soc S Afr 70:57–69

Ruxton GD, Wilkinson DM (2011) Avoidance of overheating and selection for both hair loss and bipedality in hominins. Proc Natl Acad Sci USA 108:20965–20969

Ryo M, Rillig MC (2017) Statistically reinforced machine learning for nonlinear patterns and variable interactions. Ecosphere 8:e01976

Sankaran M, Hanan NP, Scholes RJ, Ratnam J, Augustine DJ, Cade BS, Gignoux J, Higgins SI, Le Roux X, Ludwig F, Ardo J, Banyikwa F, Bronn A, Bucini G, Caylor KK, Coughenour MB, Diouf A, Ekaya W, Feral CJ, February EC, Frost PG, Hiernaux P, Hrabar H, Metzger KL, Prins HH, Ringrose S, Sea W, Tews J, Worden J, Zambatis N (2005) Determinants of woody cover in African savannahs. Nature 438:846–849

Sayre RP, Comer J, Hak C, Josse J, Bow H, Warner M, Larwanou M, Kelbessa E, Bekele T, Kehl H, Amena R, Andriamasimanana R, Ba T, Benson L, Boucher T, Brown M, Cress JJ, Dassering O, Friesen BA, Gachathi F, Houcine S, Keita M, Khamala E, Marangu D, Mokua F, Morou B, Mucina L, Mugisha S, Mwavu E, Rutherford M, Sanou P, Syampungani S, Tomor B, Vall AOM, Weghe JPV, Wangui E, Waruingi L (2013) A new map of standardized terrestrial ecosystems of Africa. Association of American Geographers, Washington, DC

Scholtz R, Kiker GA, Smit IPJ, Venter FJ (2014) Identifying drivers that influence the spatial distribution of woody vegetation in Kruger National Park, South Africa. Ecosphere 5:71

Shongwe ME, Van Oldenborgh GJ, Van Den Hurk BJJM, De Boer B, Coelho CAS, Van Aalst MK (2009) Projected changes in mean and extreme precipitation in Africa under global warming. Part I: Southern Africa. J Clim 22:3819–3837

Shongwe ME, Van Oldenborgh GJ, Van Den Hurk BJJM, Van Aalst M (2011) Projected changes in mean and extreme precipitation in Africa under global warming. Part II: East Africa. J Clim 24:3718–3733

Shorrocks B, Bates W (2015) The biology of African savannahs, 2nd edn. Oxford University Press, Oxford

Stein A, Gerstner K, Kreft H (2014) Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecol Lett 17:866–880

Torello-Raventos M, Feldpausch TR, Veenendaal E, Schrodt F, Saiz G, Domingues TF, Djagbletey G, Ford A, Kemp J, Marimon BS, Marimon BH Jr, Lenza E, Ratter JA, Maracahipes L, Sasaki D, Sonké B, Zapfack L, Taedoumg H, Villarroel D, Schwarz M, Quesada CA, Ishida FY, Nardoto GB, Affum-Baffoe K, Arroyo L, Bowman DMJS, Compaore H, Davies K, Diallo A, Fyllas NM, Gilpin M, Hien F, Johnson M, Killeen TJ, Metcalfe D, Miranda HS, Steininger M, Thomson J, Sykora K, Mougin E, Hiernaux P, Bird MI, Grace J, Lewis SL, Phillips OL, Lloyd J (2013) On the delineation of tropical vegetation types with an emphasis on forest/Savanna transitions. Plant Ecol Divers 6:101–137

Turner A, Antón M (2004) Evolving Eden. Columbia University Press, New York

Veldhuis MP, Rozen-Rechels D, le Roux E, Cromsigt JPGM, Berg MP, Olff H (2016) Determinants of patchiness of woody vegetation in an African savanna. J Veg Sci 28:93–104

Venter ZS, Cramer MD, Hawkins H-J (2018) drivers of woody plant encroachment over Africa. Nat Commun 9:2272. https://doi.org/10.1038/s41467-018-04616-8

Waldram MS, Bond WJ, Stock WD (2008) Ecological engineering by a mega-grazer: white rhino impacts on a South African savanna. Ecosystems 11:101–112

Wayman JP, Wynne RH, Scrivani JA, Reams GA (2001) Landsat TM-based forest area estimation using iterative guided spectral class rejection. Photogramm Eng Rem S 67:1155–1166

Wheeler PE (1984) The evolution of bipedality and loss of functional body hair in hominids. J Hum Evol 13:91–98

Yeaton RI (1988) Porcupines, fires and the dynamics of the tree layer of the Burkea africana savanna. J Ecol 76:1017–1029