ORIGINAL ARTICLE

The spatial distribution of vegetation types in the Serengeti ecosystem: the influence of rainfall and topographic relief on vegetation patch characteristics

D. N. Reed[1]*, T. M. Anderson[2], J. Dempewolf[3], K. Metzger[4] and S. Serneels[5]

[1]*Department of Anthropology, University of Texas at Austin, Austin, TX, USA,* [2]*Community and Conservation Ecology Group, Centre for Ecological and Evolutionary Studies, University of Groningen, Groningen, The Netherlands,* [3]*Department of Geography, University of Maryland, College Park, MD, USA,* [4]*Centre for Biodiversity Research, University of British Columbia, Vancouver, BC, Canada and* [5]*Department of Geography, Université Catholique de Louvain, Place Louis Pasteur, Louvain-la-Neuve, Belgium*

ABSTRACT

Aim The aim of this study is to introduce a structural vegetation map of the Serengeti ecosystem and, based on the map, to test the relative influences of landscape factors on the spatial heterogeneity of vegetation in the ecosystem.

Location This study was conducted in the Serengeti–Maasai Mara ecosystem in northern Tanzania and southern Kenya, between 34° and 36° E longitude, and 1° and 2° S latitude.

Methods The vegetation map was produced from satellite imagery using data from over 800 ground-truthing points. Spatial characteristics of the vegetation were analysed in the resulting map using the FRAGSTATS software package. Average patch area and nearest neighbour distance (NND) were determined for grassland, shrubland and woodland vegetation types. The heterogeneity of vegetation types was estimated with Simpson's diversity index (D). Structural equation modelling (SEM) was used to explore the relationships between the spatial characteristics of vegetation and three predictor variables: annual rainfall, coefficient of variation (CV) in annual rainfall, and topographic moisture index (TMI).

Results A vegetation map is presented along with a detailed summary of the distribution of land-cover classes and spatial heterogeneity in the ecosystem. Significant relationships were found between vegetation diversity (D) and TMI, and also between D and average rainfall. The average area of grassland patches showed significant relationships with average rainfall, with rainfall CV and with TMI. Grassland NND was positively associated with average rainfall. Woodland patch area showed a unimodal response to average rainfall and a negative linear association with TMI. Woodland NND showed a U-shaped association with annual rainfall and a weaker positive linear association with TMI. An acceptable model that explained variation in shrubland patch characteristics could not be identified.

Main conclusions The vegetation map and analysis thereof resulted in three significant causal explanatory models that demonstrate that both rainfall and topography are important contributors to the distribution of woodlands and grasslands in the Serengeti. These findings further indicate that changes in patch characteristics have a complex interaction with rainfall and with topography. Our results are concordant with recent studies suggesting that percent woody cover in African savannas receiving less than *c.* 650 mm year^{-1} is bounded by average annual rainfall.

Keywords

African savanna, fragstats, fuzzy classification, fuzzy convolution, physiognomic vegetation map, SEM, vegetation heterogeneity.

*Correspondence: Denné Reed, Department of Anthropology, University of Texas at Austin, 1 University Station C3200, Austin, TX 78712, USA.
E-mail: reedd@mail.utexas.edu

INTRODUCTION

As a World Heritage Site and home to some of Earth's largest free-ranging ungulate herds, the Serengeti–Maasai Mara ecosystem (hereafter simply the Serengeti) is a treasure for biodiversity conservation and one of the best-studied natural laboratories in the palaeotropics (Sinclair, 1995; Sinclair *et al.*, 2008). Vegetation heterogeneity provides the physical template on which the most conspicuous and important ecosystem processes in the Serengeti operate (Anderson *et al.*, 2008), including the annual migration of nearly two million wildebeests and zebras (Thirgood *et al.*, 2004). Moreover, the spatial distribution of Serengeti vegetation types is key to understanding animal behaviour (Fryxell *et al.*, 2005; Hopcraft *et al.*, 2005; Packer *et al.*, 2005), biogeochemistry (McNaughton *et al.*, 1997; Anderson *et al.*, 2004), hydrology (Wolanski & Gereta, 2001; Gereta *et al.*, 2004) and human impacts on ecosystems (Scholes & Archer, 1997; Serneels & Lambin, 2001; House *et al.*, 2003; Lamprey & Reid, 2004; Thirgood *et al.*, 2004). In particular, the distribution of woody vs. herbaceous vegetation and the density, cover and height of that vegetation are central to understanding general functions and processes of savanna ecosystems (Solbrig, 1996; Sankaran *et al.*, 2004).

No ground-truthed, spatially explicit description of vegetation structure or heterogeneity has yet been produced for the Serengeti, despite the importance of vegetation patterns for a general understanding of the ecosystem and over sixty years of active research in the national park and associated protected areas. Furthermore, except for anecdotal descriptions relating to maps of soil (deWit, 1978; Jager, 1982) or to dominant tree species identity (Herlocker, 1974), an understanding of what factors control the spatial variation of vegetation types at regional spatial scales in the Serengeti is lacking. At the scale of the African continent, climate and topography emerge as key controllers of woody vegetation cover (Sankaran *et al.*, 2004, 2005) and woody plant species richness (O'Brien *et al.*, 2000; Whittaker *et al.*, 2001; Field *et al.*, 2005). Research on vegetation community composition at finer spatial scales (Coughenour & Ellis, 1993; Witkowski & O'Connor, 1996; Burnett *et al.*, 1998; Nichols *et al.*, 1998; Urban & Keitt, 2001) is consistent with the idea that climate and topography are controlling factors for the distribution of vegetation types across landscapes. However, the balance of factors controlling vegetation structural heterogeneity, including the size, diversity and isolation of vegetation patches, is not well understood in savanna ecosystems.

This study seeks to: (1) describe the spatial patterns of vegetation type and structure for one of the last fully intact savanna-grassland ecosystems on Earth, the Serengeti–Maasai Mara, and (2) quantify the relative significance of rainfall and topography on vegetation heterogeneity. Vegetation heterogeneity in this context refers to the quantitative description of patch characteristics of different vegetation types. Patch characteristics such as size diversity and isolation are key properties that organisms recognize and respond to as they navigate and interact with their environment (O'Neill *et al.*, 1988; Kolasa &

Waltho, 1998; Ritchie, 1998, 2002). The description of Serengeti vegetation heterogeneity will provide a baseline for conservation measures and for an understanding of the ecosystem extending far into the future. We conducted detailed vegetation mapping throughout the Serengeti and combined the results with remotely sensed data to produce a detailed structural vegetation map. Building on this map, we used structural equation modelling (SEM) (Grace, 2006) to quantify the relative contributions of mean annual rainfall, annual variation in rainfall and topographic variation to patterns of vegetation heterogeneity. Beyond clarifying internal controls over vegetation heterogeneity, a description of the relationship between rainfall, topography and vegetation pattern may elucidate the factors controlling vegetation dynamics in other savanna ecosystems (House *et al.*, 2003; Sankaran *et al.*, 2004).

BACKGROUND

The earliest cartographic descriptions of the ecosystem come from intensive scientific exploration during the 1970s. Herlocker (1974) and Herlocker *et al.* (1993) prepared a woody plant species map of the Serengeti National Park. This work was an extension of earlier efforts to map regions of similar land-cover types, such as the zonal land classification scheme of Gerresheim (1971). Two soil maps followed these efforts, one for the Serengeti plains (deWit & Jeronimus, 1977) and another for the woody regions of the western corridor and the northern extension (Jager, 1982). Climate patterns and precipitation were also described in detail at this time (Norton-Griffiths, 1975). These early maps benefited from extensive ground-truthing and were mapped using aerial photographs, but map features were drawn by hand and this limited the detail and accuracy with which complex land-cover boundaries were recorded.

The advent of GIS and satellite remote sensing brought about several commercial land-cover classifications, such as those by Hunting Services and the Tanzania Natural Resources Information Centre (TANRIC). These maps, however, were produced using unsupervised interpretation of satellite imagery without extensive ground-truthing. The structural vegetation map that we present in this paper is the first detailed map of the Serengeti ecosystem based on extensive ground-truthing and benefiting from the spatial and spectral precision afforded by satellite remote sensing.

Ground-truthing was conducted in the three core protected areas: the Ngorongoro Conservation Area (NCA), the Serengeti National Park and the Maasai Mara National Reserve, and in the northerly adjacent Maasai group ranch Koyaki. The map covers regions surrounding these core areas for which the classification signatures should maintain reasonable accuracy.

MATERIALS AND METHODS

The Serengeti–Maasai Mara ecosystem straddles the Tanzania–Kenya border in East Africa between 34° and 36° E longitude, and 1° and 2° S latitude. The Serengeti National Park in

Tanzania encompasses an area of *c.* 14,800 km², but the larger ecosystem – defined as the area covered by the wildebeest migration – extends into the neighbouring Maasai Mara National Reserve and the adjacent Narok region to the north in Kenya, Ngorongoro Conservation Area to the south-east, the Loliondo Game Controlled Area to the east, Maswa Game Reserve to the south-west, and the Ikoronogo and Grumeti Game Controlled Areas to the north-west. In total, the ecosystem covers an expanse of roughly 24,000 km², as shown in Fig. 1.

Landsat ETM+ data

Four Landsat 7 ETM+ scenes served as the principal data source: Path 169 Row 61 collected 2 February 2000, Path 169 Row 62 collected 2 February 2000, Path 170 Row 61 collected 30 October 1999 and Path 170 Row 62 collected 30 October 1999. The supervised classification was conducted in the ERDAS Imagine software package (v. 8.7, Leica Geosystem, Norcross, GA, USA) using a fuzzy-classification approach with post-classification fuzzy convolution. The study area was divided into woodland, grassland and western corridor zones. Each zone was classified separately and later merged together. Classification signatures were evaluated iteratively for separability (the degree to which spectral signatures do not overlap one another) and exhaustibility (the degree to which all signatures combined cover the total variation in the source imagery) and then applied to the image using a fuzzy-classification algorithm that stored three membership grades per pixel. A fuzzy convolution filter was used to reduce the speckling of the classification, and clumps of < 9 pixels were removed so that the minimum mappable unit is *c.* 8100 m², or slightly < 1 ha.

The Landsat images are largely cloud- and haze-free. Small clouds dot the eastern margins of path 169. Both scenes for path 170 were collected by the satellite on the same day (30 October 1999), as were both scenes for path 169 (2 February 2000). The scenes for each path were stitched together and then orthorectified to produce two path-mosaic images. The path 170 mosaic was rectified using 21 ground control points and a total root mean square (RMS) error of 36.4 m. The path 169 mosaic used 84 points with an RMS error of 40.5 m. The path mosaics were georectified using a nearest-neighbour algorithm in order to preserve original pixel values.

The ETM+ sensor aboard Landsat 7 records data from seven spectral bands, each at *c.* 30-m spatial resolution (although band 6 is coarser) and in addition these is one pan-chromatic band at *c.* 15-m spatial resolution (Irish, 2001; Lillesand *et al.*, 2008). For each scene, bands 1–5 and 7 were combined into a layer stack. Band 6 primarily represents emitted thermal radiation, and it was excluded owing to its lower spatial resolution and low performance in signature separability measures. Five additional layers were calculated and added to each scene to create an 11-band image stack: the first three principal components of each scene, a texture layer calculated from the variance in the first principal component, and a normalized difference vegetation index (NDVI) layer. Image texture incorporates information about neighbouring pixels into the classification and aids in distinguishing dense grassland from dense grassed woodland. Woodland areas generally have greater texture whereas grasslands are much more homogeneous.

Land-cover classification scheme

The land-cover classes portrayed in the final map are based on the system developed by Grunblatt *et al.* (1989). This system is hierarchical in that different levels of precision can be applied depending on the data available, and more specific classes can be telescoped inside broader classes and
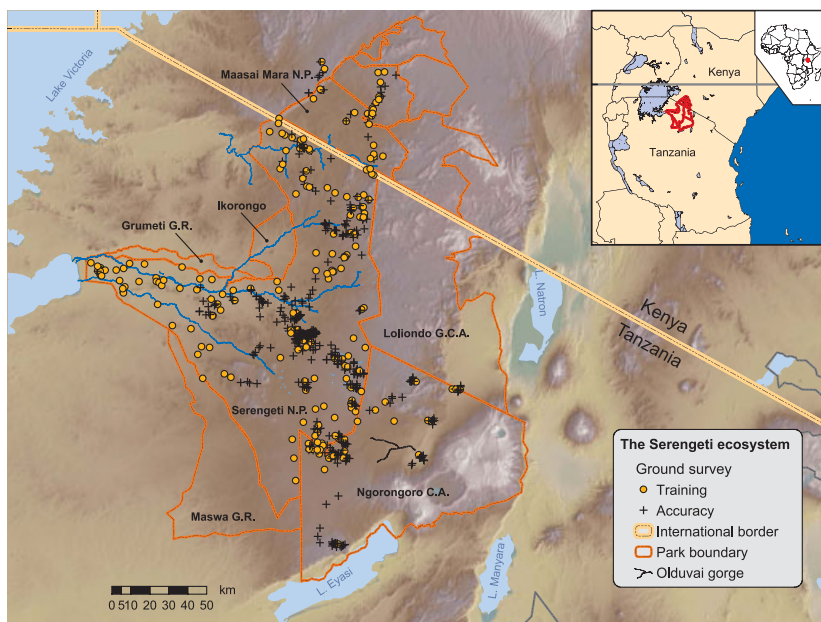

Figure 1 Shaded relief map of the study area showing international borders and protected area boundaries. Points indicate ground-truthing locations. Crosses indicate points used for classification and circles show those used for accuracy assessment and validation. Illumination for hill-shading comes from the south-east.

Table 1 Summary of terms used in the land-cover classification scheme.

Term	Symbol	Description
Life-forms		
Forest	F	Single-stem woody plants generally taller than 1.5 m and at density > 50% canopy cover
Woodland	W	Single-stem woody plants generally taller than 1.5 m at densities < 50% canopy cover
Shrubland	S	Multi-stem woody plants generally below 2 m
Grassland	G	Herbaceous vegetation generally below 2 m including grasses, sedges, etc.
Bare	B	Unvegetated patches including rock, and bare earth
Density modifiers		
Closed	c	80–100% canopy cover
Dense	d	50–79% canopy cover
Open	o	20–49% canopy cover
Sparse	s	2–19% canopy cover

See text for details and for examples of the vegetation classification scheme.

vice versa (Table 1). The hierarchy has four levels. At the broadest level (Level 1), the vegetation is described in terms of the primary life-form and its density, for example open grassland (oG), or dense forest (dF). In general, candidates for the primary life-form must have a percentage canopy cover > 20%, and, within those, preference is given to trees over shrubs, to shrubs over grass, and to grass over bare ground. The Level 1 adjective describes the primary life-forms' canopy density.

For example, a plot with 25% trees, 15% shrubs, 30% grass and 70% bare ground is an open woodland (oW). If an area has no life-form with a percentage canopy cover > 20% but some with a cover > 2% then the life-form with the greatest canopy cover is used along with the sparse modifier; for example, a canopy coverage of 0% trees, 2% shrubs, 10% grass and 90% bare ground would be classified as sparse grassland (sG, not to be confused with shrubbed grassland SG – see below).

Level 2 designations include a secondary life-form as a descriptive modifier to the primary life-form. The terms for the secondary life-form are similar to those for the primary: Treed, Shrubbed and Grassed. The Level 2 life-form may be used when a form other than the primary form attains 20% density, with preference following the same order as earlier (trees, shrubs, grass). The plot described in the first example above (T25%, S15%, G30%, B70%) would have a Level 2 classification of open grassed woodland (oGW), whereas another with slightly greater shrub cover (for example T25%, S22%, G30%, B70%) would be an open shrubbed woodland (oSW). Both the density adjective and the secondary life-form are descriptors of the primary life-form. The system of Grunblatt *et al.* (1989) gives emphasis to wooded and shrubbed categories, allowing them to be included as secondary descriptors if they are present at levels between 2% and 19%

when no other life-form is present as a true secondary candidate (i.e. > 20%). To illustrate: if a plot has T5%, S15%, G70%, B30% it would be a dense treed grassland (dTG). Further examples and details about the system are available in Grunblatt *et al.* (1989).

Ground-truthing

Ground-truthing began in 1998 and ended in 2002. In total, 859 ground-truthing points were sampled using three methodologies. Plots using the first method, modified-Whittaker plots (Stohlgren *et al.*, 1995), are intensive surveys conducted on 1000 m^2 grassland plots in which the hierarchical structure of the plant species community is examined through overlapping plots of different sizes. Modified-Whittaker plots include data on species composition in addition to physiognomic structure. Plant specimens from modified-Whittaker plots were identified using the herbarium collections at the Mweka Wildlife Research College and the Serengeti Wildlife Research Institute, Seronera. A total of 133 modified-Whittaker plots are in the data set. Another 376 ground-truthing points were conducted using a second method, namely 9-point plots. Like the modified-Whittaker plots, the 9-point plots have a nested structure. The main sampling area is 100 × 100 m (10,000 m^2 = 1 ha); this plot is subdivided into nine subplots with measurements taken from a circular region 8 m in diameter at the centre of each subplot. Vegetation measurements were taken in the 8-m-radius area, including tree and shrub species identity, height, and crown diameter. Measurements were then averaged over the entire plot (100 × 100 m) area and standardized to conform to the Grunblatt convention.

Geographical placement of the modified-Whittaker and 9-point plots was stratified-random. General locations were chosen from randomly generated geographic coordinates, but the specific placement of the plot was sometimes altered to avoid local hazards.

A simpler methodology was used for the remaining 350 ground-truthing points. A 30 × 30 m area was subdivided into four quadrats, and the percentage canopy covers of trees and shrubs were estimated visually by projecting canopy edges to the ground at midday and estimating the percentage area shaded in each 15 × 15 m quadrat. Estimates were averaged over all quadrats to calculate the summary value for each 30 × 30 m plot. These data were collected by two teams, with cross-training and comparison to reduce inter-observer error between teams. Grass density was measured by visual inspection of a fixed-area 0.5-m^2 metal rectangle (Daubenmire plot) randomly placed throughout the 30-m plot. Bare ground was taken as the converse of the grass density (1 – grass). Notes on species composition along with Hi8 video and digital photographs supplemented these data.

The initial placement of vegetation plots was stratified-random, and often conducted in transects. Later plots were chosen interactively by consulting the satellite imagery and preliminary iterations of the supervised classification created in the field to target areas not yet classified. More detailed

accounts of the data collection methodology are available in Reed (2003) and Anderson *et al.* (2007).

Signature development

Training areas were developed by semi-automated region-growing centred at 174 of the 859 ground-truth survey points. The ERDAS Imagine software grows an area of interest (AOI) starting from the centre point and appending adjacent pixels that are spectrally similar. Regions were grown until no sufficiently similar adjoining pixels could be found or until an area of *c.* 100 pixels was sampled. The threshold Euclidean spectral distance for growing regions was set interactively by the operator and depended on the type of signature being developed. Signatures were tested for separability using a transformed divergence measure, and signatures with low separability were merged and retained if the merged result was a normally distributed set of training pixels. Otherwise, the signatures were kept separate. In this way, several subclasses were developed for each ground cover type. Preliminary classifications were run to test for feature space exhaustibility (i.e. the completeness with which the signature set covered the hyper-dimensional space of the imagery), and new signatures developed as needed.

Classification

Ecotonal boundaries are often gradual or arranged in complex patterns that follow edaphic conditions, creating a mosaic of land-cover types over several spatial scales. Moreover, homogeneous land-cover pixels are rare, and differences between land-cover classes are often subtle. Fuzzy classification copes better with pixels of mixed makeup and complex spatial disposition than does standard maximum-likelihood classification (Leica, 2006). The general routine of fuzzy classification is very similar to that of maximum likelihood. However, whereas maximum-likelihood classification assigns each pixel to a single class, fuzzy classification assigns multiple classes to each pixel. In this study, each pixel was assigned a grade for the three closest classes, returning a three-band image with the three bands containing the primary, secondary and tertiary class assignments respectively. A related, three-band distance file records the class grades for each pixel as well, so a record is maintained of class assignment along with a statistical measure of how well that pixel is classified to the different classes.

Fuzzy convolution is a contextual filter that reduces the speckling common to maximum-likelihood output. Fuzzy convolution filtering moves pixel-by-pixel through the image and at each pixel examines neighbouring pixels in an $N \times N$ window surrounding the focal pixel and across each band. We employed an $N = 3$ convolution window across the three-band fuzzy-classification, resulting in a $3 \times 3 \times 3$ convolution cube. The class assignment of the centre pixel is evaluated in light of the surrounding pixels, and weighted by their spatial distance to the focal pixel and the grade values (or spectral distance) of their respective classes. The class assignment of the focal pixel

may be changed if there are enough, well-classified surrounding pixels. Finally, the CLUMP and ELIMINATE algorithms in ERDAS Imagine were used to remove clusters of land cover smaller than 9 pixels in size, which is equal to an area of 8100 m^2, or nearly 1 ha.

Methods for the analysis of vegetation patches

To understand how rainfall and topographic factors influence the distribution of vegetation across the Serengeti ecosystem, relationships among rainfall, topographic variation and vegetation patch characteristics were analysed. Three vegetation characteristics were calculated from the Serengeti vegetation map using version 3.0 of the freeware program FRAGSTATS (McGarigal *et al.*, 2002): patch area, patch Euclidean nearest-neighbour distance (NND) and patch diversity. Patch area and NND were determined separately for three broad vegetation types, namely grassland, shrubland and woodland (Level 1 classification), whereas patch diversity was determined by the heterogeneity of all three vegetation types using Simpson's index (*D*). Further description of the metrics used and of the associated equations can be found at the FRAGSTATS website, http://www.umass.edu/landeco/research/fragstats/fragstats.html.

Two rainfall variables, average annual rainfall and the coefficient of variation (CV) in annual rainfall, and one landscape variable, topographic moisture index (Beven & Kirkby, 1979; Mackey, 1994; Wilson & Gallant, 2000), were used as independent predictors to explain variation in vegetation characteristics. Precipitation estimates were generated using 40 years of monthly rainfall data collected from 58 rain-gauges across the ecosystem (Serengeti Ecological Monitoring Program). A computer program, PPTMAP (Coughenour, 2006), was employed to create average monthly and mean annual precipitation estimates for the study region. PPTMAP uses available precipitation data from multiple weather stations, and spatially interpolates the data to develop a grid-cell map of precipitation across the region. The interpolation technique used was inverse distance weighting, corrected for significant effects of elevation. A regression equation was developed within the program, relating precipitation to elevation, based upon the station data. The slope of the regression line of elevation and precipitation provided a correction of millimetres of rainfall per metre elevation difference between any location and any observation station. Precipitation was modelled at 1 km × 1 km resolution (Fig. 2 insert).

Topography affects the distribution of vegetation by altering water runoff and hence the hydrological condition of the soil. The topographic moisture index (TMI) was used to characterize the influence of topographic variation on the spatial variation of soil water content, where TMI = $\ln(As/\tan(sl))$, where *As* is the specific catchment area (the upslope area contribution to runoff divided by the grid-cell side length), and *sl* is the local slope in degrees. The value of the index increases in flat, low-lying areas and decreases as landscapes become steeper and lose more moisture to runoff. TMI was calculated using the freeware software package Terrain Analysis
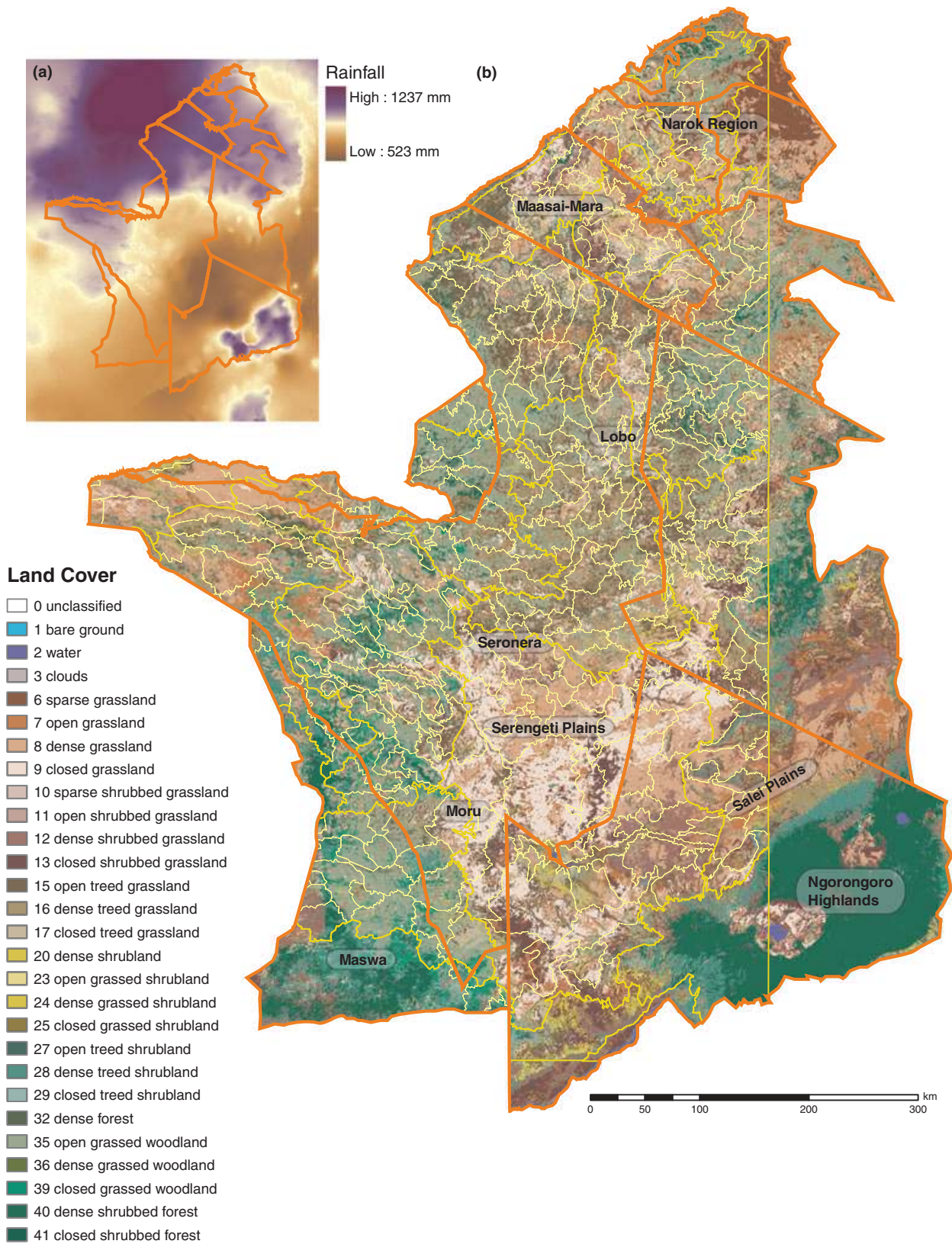
Rainfall

High : 1237 mm

Low : 523 mm

(a)

(b)

Narok Region

Maasai-Mara

Lobo

Seronera

Serengeti Plains

Saler Plains

Moru

Ngorongoro Highlands

Maswa

Land Cover

- 0 unclassified
- 1 bare ground
- 2 water
- 3 clouds
- 6 sparse grassland
- 7 open grassland
- 8 dense grassland
- 9 closed grassland
- 10 sparse shrubbed grassland
- 11 open shrubbed grassland
- 12 dense shrubbed grassland
- 13 closed shrubbed grassland
- 15 open treed grassland
- 16 dense treed grassland
- 17 closed treed grassland
- 20 dense shrubland
- 23 open grassed shrubland
- 24 dense grassed shrubland
- 25 closed grassed shrubland
- 27 open treed shrubland
- 28 dense treed shrubland
- 29 closed treed shrubland
- 32 dense forest
- 35 open grassed woodland
- 36 dense grassed woodland
- 39 closed grassed woodland
- 40 dense shrubbed forest
- 41 closed shrubbed forest

km

0   50   100   200   300

Figure 2 Serengeti–Maasai Mara, showing (a) rainfall map insert, illustrating the pronounced rainfall gradient in the Serengeti, and (b) structural vegetation map of the Serengeti–Maasai Mara ecosystem with protected area boundaries and Gerresheim (1971) land-cover subregions outlined in yellow. Heavy lines delimit regions, lighter lines delimit subregions. The vertical line along the right is the eastern border of Gerresheim's map.

System 2.0.9 from the hydrologically conditioned version of the Shuttle Radar Topography Mission (SRTM) digital elevation model (Lehner *et al.*, 2006).

The landscape boundaries used to calculate rainfall, topography and vegetation metrics were based on the Serengeti landscape classification of Gerresheim (1971). Gerresheim's classification dissected the Serengeti ecosystem into distinct 'land regions' and 'land subregions' representing main landscape types that share a common geological history, comparable geomorphic and topographic influence, and similar climate. The relationships between vegetation patch metrics and the variables representing rainfall and topography were analysed by averaging the variables within Gerresheim land-classification subregions, which range in size from 1 to 500 km^2 (median area = 49.2 km^2). For our analyses, 430 subregions, nested inside 14 land regions occurring in protected areas (of Gerresheim's original 20), were used because: (1) our ground-truthing overlapped generously with those land regions and thus our confidence in the analysis was high; (2) vegetation patch characteristics outside the protected areas were subject to human influences and other disturbances for which we had no reliable measures; and (3) a few very small subregions (< 1 km^2) were discarded from the analysis. For the subregions included in the analysis, rainfall, CV rainfall and TMI for each land region were the arithmetic averages of all grid cells in each subregion. Likewise, vegetation patch area and NND were the average values calculated from all patches within a vegetation type occurring within each Gerresheim subregion. Vegetation diversity, measured by Simpson's index (*D*), was calculated using all pixels within a subregion; as such, *D* is the probability of randomly drawing two pixels of the same vegetation type from a particular subregion.

Relationships among independent and response variables were analysed by means of structural equation modelling (SEM) in the software package AMOS 16.0 (Arbuckle, 2007). The *a priori* model included all possible effects of rainfall, CV rainfall and TMI on vegetation characteristics. Direct interactions among variables (single-headed arrows in resulting figures) were estimated as standardized coefficients from the covariance matrix. Non-directional standardized correlation coefficients were also calculated among independent and among observed variables (double-headed curved arrows in resulting figures). In the case of significant nonlinear relationships between independent and response variables, polynomial relationships were modelled by constructing composite variables according to Grace & Bollen (2008). A final model was obtained by eliminating insignificant unidirectional paths until optimum model fit was achieved. The final model fit was evaluated with an overall chi-square statistic for model goodness-of-fit: a large *P*-value suggests that the true model does not differ from the observed model result. Exploratory analyses confirmed that subregion size was uncorrelated with predictors and response variables, and therefore it was not necessary to correct for this effect in the model. The response variables patch area and NND were log-transformed to stabilize the covariance matrix during the analysis.

RESULTS

The results of the classification are presented in the map shown in Fig. 2. Digital versions of the map are available from the first author.

Accuracy assessment

Ground-truth points not used for signature development (*n* = 695) were used for accuracy assessment (Fig. 1). Accuracy was measured for those classes with more than 20 accuracy assessment points, and the results are presented in Table 2. Accuracy results are presented in two ways. Upper values in the table show the discrete matches for each class (i.e. the number of times that the code of the accuracy assessment point matched that of the classified image). This is the standard method of accuracy assessment. A second, fuzzy value is also calculated, where reference points are considered to match the map if they hit the identical or an adjacent category (i.e. a category that can be reached from the current category by simply increasing or decreasing the percentage cover of one vegetation type by one class). To illustrate: if an open grassland (oG) accuracy assessment point fell onto a pixel that was classified as dense grassland (dG), sparse grassland (sG) or open treed grassland (oTG), it would be coded as a match rather than as an error. However, if the same point fell on a pixel coded as closed grassland (cG) or as dense treed grassland (dTG), this would be counted as an error. This method of assessing the accuracy of maps is described in greater detail by Jensen (1996).

The results given here present a conservative assessment of accuracy, as many of the spectrally most distinct classes of ground cover, such as water, bare ground and dense forest, have few ground-truth accuracy assessment points. This reflects data collection priorities that focused on grassland and wooded grassland areas.

Classifying shrubs (low woody vegetation) is particularly difficult, as indicated by the two classes that performed worst, dSG and cSG. However, the inaccuracies are consistent, such that misclassified pixels end up in categories of similar ecological composition: open grassland most often gets mistaken for dense grassland and less often for open grassed woodlands. For this reason the fuzzy accuracy results produce a marked improvement.

Results of the spatial analysis of vegetation patches

Vegetation diversity

The final model predicting Simpson's index (*D*) suggested that vegetation diversity is influenced by variation in both rainfall and topographic factors (Fig. 3; χ^2 = 4.6, d.f. = 3, *P* = 0.21; r^2 = 0.27). The strongest standardized effect was a nonlinear (unimodal) relationship with average annual rainfall. In bivariate space, peak vegetation diversity occurred at 905 mm year^{-1}. After accounting for the relatively strong correlation between annual rainfall and CV annual rainfall

Table 2 Accuracy assessment matrix, with rows showing pixels as they appear on the vegetation map compared against columns showing the vegetation recorded during ground surveys at the same point.

	Reference data								Row total
Veg. class	oG	dG	cG	dSG	cSG	oTG	dTG	oGW	
Open grassland (oG)	8	6	2	0	2	2	1	1	22
Dense grassland (dG)	6	26	6	5	5	4	3	1	56
Closed grassland (cG)	4	20	31	2	6	2	7	1	73
Dense shrubbed grassland (dSG)	3	4	2	3	4	1	3	0	20
Closed shrubbed grassland (cSG)	0	5	12	6	11	0	0	0	34
Open treed grassland (oTG)	2	7	2	4	7	21	12	3	58
Dense treed grassland (dTG)	1	1	1	1	0	5	11	4	24
Open grassed woodland (oGW)	3	5	0	2	1	7	11	12	41
Column total	27	74	56	23	36	42	48	22	
Column statistics									
Column correct	8	26	31	3	11	21	11	12	
Fuzzy correct	16	57	49	15	21	35	40	19	
User's accuracy (%)	30	35	55	13	31	50	23	55	
Fuzzy user's accuracy (%)	59	77	88	65	58	83	83	86	
Row statistics									
Row correct	8	26	31	3	11	21	11	12	
Fuzzy correct	16	46	57	14	29	38	22	30	
Producer's accuracy (%)	36	46	42	15	32	36	46	29	
Fuzzy producer's accuracy (%)	73	82	78	70	85	66	92	73	
Overall accuracy (%)	38								
Overall fuzzy accuracy (%)	77								

The leftmost label "Classification data" runs vertically alongside the veg. class rows.

Grey boxes indicate classes that are similar to one another and combined in the fuzzy assessment. See Table 1 for a description of the vegetation classes.

(Fig. 3; $r = -0.65$), there remained no significant effect of CV rainfall on vegetation diversity at $\alpha = 0.05$. Finally, the negative path between TMI and D indicates that vegetation diversity, and thus the coexistence of major plant functional types, increases with topographic variation in rugged, more dissected areas of the Serengeti ecosystem.

Patch characteristics by vegetation type

The final model predicting grassland patch characteristics included significant paths from all independent variables to patch area and one significant path to NND from annual rainfall (Fig. 4a; $\chi^2 = 3.7$, d.f. = 3, $P = 0.30$). The strongest standardized effect was a negative association between grassland patch area and annual rainfall; together with this effect, positive relationships between CV rainfall and TMI explained 25% of the variation in grassland patch area (Fig. 4a). Although annual rainfall also had a positive effect on NND, the amount of variation explained by this effect was only 3%.

The final model predicting woodland patch characteristics included nonlinear effects of annual rainfall and linear effects of TMI on patch area and NND (Fig. 4b; $\chi^2 = 6.7$, d.f. = 5, $P = 0.24$). The nonlinear relationship between annual rainfall and NND was U-shaped and corresponded to a bivariate minimum NND at an annual rainfall of 887 mm year^{-1}. For woodland patch area, the relationship was unimodal with rainfall and displayed a peak in bivariate space at 852 mm year^{-1}. Unlike grasslands, TMI was negatively related to woodland patch area and positively associated with NND.

For the shrubland vegetation type, once paths above $\alpha = 0.05$ were removed, the final model predicting patch
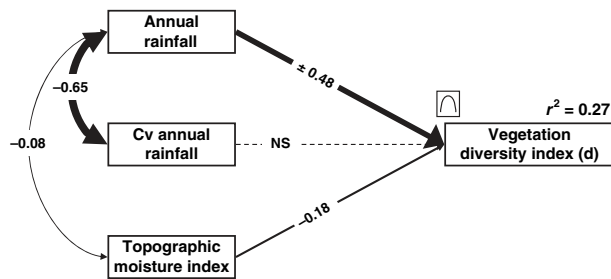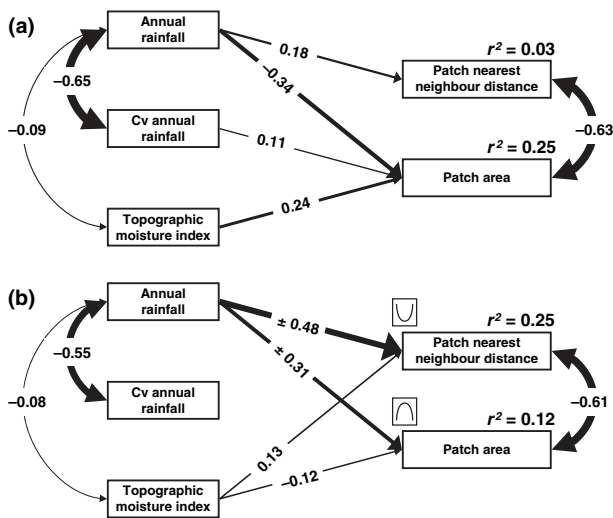

Figure 3 Graphical results of the structural equation model testing relationships among annual rainfall, CV annual rainfall, topographic moisture index and vegetation diversity (Simpson's index, D). Single-headed straight arrows indicate directional relationships between independent and response variables, and curved arrows represent correlations between independent variables. Values associated with unidirectional arrows represent standardized path coefficients, and those associated with curved arrows are correlation coefficients (line weights are proportional to the strength of the coefficients). The $\pm$ symbol represents a nonlinear relationship that was modelled with a composite variable according to Grace & Bollen (2008), and the small box indicates the nature of the nonlinear relationship (either unimodal or U-shaped). The dashed line signifies a non-significant (n.s.) path at $\alpha = 0.05$. The final model had a significant fit (see Results) and explained *c*. 27% of the variance in the response variable.

Figure 4 Graphical results of the structural equation model testing relationships among independent variables and patch characteristics for (a) grassland and (b) woodland vegetation types. Meanings of arrows and symbols are as in Fig. 3. Only significant paths at α = 0.05 are shown. See the Results section for further explanation of the model results.

characteristics displayed a poor fit and was deemed statistically unacceptable. The exploration of possible nonlinear relationships also yielded no acceptable model to explain shrubland patch characteristics.

DISCUSSION

Savanna ecosystems are defined as having a continuous grass understorey in conjunction with variable amounts of woody vegetation (Talbot & Kesel, 1975; Cole, 1986). Grassland is the most abundant land-cover class in the Serengeti (*c.* 61% of the land cover in the ecosystem). The grassland plains situated in the southern Serengeti National Park and northern portions of the Ngorongoro Conservation Area make up roughly 30% of the total map area and contain the largest continuous patches of homogenous land cover. The homogeneity in land cover belies a heterogeneity in plant species beta diversity (Anderson *et al.*, 2008) and nutrient properties, such as leaf nutrients and above-ground biomass (Anderson *et al.*, 2007) germane to herbivores.

The Serengeti plains are spatially coincident with local minima in mean annual rainfall caused by rain shadows cast from the Ngorongoro highlands to the south and the Gol mountains to the east, with the driest point falling in the Ol'Balbal depression near Olduvai Gorge (Fig. 2). Moving north and west, precipitation gradually increases, as does the incidence of woody land cover as seen in the map (Fig. 2). Near the park headquarters at Seronera there is a sudden transition to treed grasslands and grassed woodlands. The transition is marked first by the presence of trees along drainages and soon after by stands of trees that intrude into the grassland matrix to produce a mosaic of woodland and

grassland. Similarly, to the west grasslands abruptly give way to woodlands along hill slopes near Moru. These transitions, clearly visible in the vegetation map, correspond with the Gerresheim (1971) boundaries for region 11 (grassland plains), indicating that this landscape transition has been stable for at least the past 25 years. Recent fossil evidence suggests that open grasslands have been an important component of the ecosystem since the start of the Pleistocene (Bamford *et al.*, 2008).

The SEM analysis of patch characteristics indicates a relatively strong link between mean annual precipitation and the patch characteristics considered: patch diversity, inter-patch distances (NND) and patch area (Fig. 4). The increase in grassland patch area with rainfall is explained in part by the large contiguous areas of grassland that dominate in the drier portions of the rainfall gradient. The grassland plains form a matrix of grass in which woody vegetation is dispersed, generally along riparian margins. Grassland patches in the arid portions of the ecosystem are contiguous, which explains the relatively low NND of grassland patches and the low land-cover diversity at low rainfall. In contrast, woodland patches are widely separated in the grassland matrix at low rainfall but increase in number and density at the middle of the rainfall gradient (*c.* 887 mm year^{-1}). This pattern also explains the change in vegetation diversity (*D*), which peaks at a similar point along the rainfall gradient (*c.* 900 mm year^{-1}).

The SEM analysis also recovered a weaker but significant link between the topographic moisture index and patch characteristics, even after having controlled for the effects of rainfall, including greater landscape diversity in regions of greater relief compared with relatively flat areas dominated by grassland. This result implies that topographic heterogeneity in the Serengeti contributes directly to the coexistence of woody and herbaceous vegetation patches across the landscape. These results are consistent with the idea of Coughenour & Ellis (1993), that vegetation structure is hierarchically constrained by physical factors and that topographic effects on water redistribution and availability are important determinants of vegetation structure in tropical savannas.

The patterns of structural vegetation modelled in the SEM raise the following two important and long-standing questions: (1) What factors are responsible for the large, uninterrupted grassland plains? (2) Why do grasslands and woodland coexist in a complex mosaic in the areas beyond the grassland plains? Given the coincidence of large grassland patches within regions of minimum rainfall, it is reasonable to suggest, as many have, that rainfall has a causal effect on vegetation structure, especially on the exclusion of trees and the establishment of grassland plains, by limiting moisture availability. However, the unimodal relationship of mean annual rainfall with landscape diversity and woody patch area suggests that the relationship between rainfall and vegetation structure is complex. Indeed, several biotic and abiotic drivers have been suggested to influence the structure of tropical savannas, including, but not limited to, mean annual rainfall, annual variability of rainfall, fire regime, evapotranspiration, altitude,

topography, soil texture, soil nutrient availability and herbivory (Belsky, 1990; Scholes & Archer, 1997).

Taken together, topography, rainfall and rainfall variability explain 25% of the variance in grassland patch area, suggesting that the model could be improved in future analyses by including the factors listed above or by sampling at a finer spatial scale. At low rainfall, soil factors especially may play a critical role. The two most comprehensive studies of Serengeti soils highlight the importance of soil origin and of their physical and structural properties as determinants of vegetation structure (deWit, 1978; Jager, 1982). In general, soils derived from a volcanic origin overlying a shallow petrocalcic layer (hard-pan), such as those in the Serengeti plains, support grasslands, whereas soils derived from Pre-Cambrian basement rock support woodlands (Dawson, 1964; Hay, 1976; deWit, 1978; Belsky, 1990, 1995). Within the woodland areas north and west of the Serengeti plains, soil moisture retention characteristics were strongly associated with the dominant vegetation type (Jager, 1982). Specifically, soils with water infiltration rates > 2.5 cm h^{-1} and a depth of penetration of > 50 cm are generally overlain by woodland (Jager, 1982). Thus, soil origin, depth of the petrocalcic horizon and water retention properties are likely candidates for the maintenance of continuous grasslands on the plains and for the relatively sharp transition from grassland to woodland in the vicinity of Seronera.

A recent study of African savannas by Sankaran et al. (2005) argued that the maximum density of woody vegetation in savannas experiencing < 650 mm $year^{-1}$ of rainfall is limited by annual rainfall. These environments they term 'climatically determined savannas', in contrast to the 'disturbance controlled savannas' that preside at higher rainfall levels. Above c. 650 mm $year^{-1}$ there is adequate moisture available such that woody vegetation should form a closed canopy and exclude open grasslands. That this does not happen in many areas is, they argue, a result of disturbance intensity from fire, large herbivores and similar factors. For comparative purposes, following Sankaran et al. (2005), we plotted the percentage of each Gerresheim subregion classified as woodland or shrubland against mean annual rainfall and analysed the plot using bent-cable quantile regression with the QUANTREG package in R (Sankaran et al., 2005; Chiu et al., 2006; Koenker, 2008). The results of the analysis, as illustrated in Fig. 5, are consistent with the model of Sankaran et al. (2005): the percentage of the landscape in a woody state across the Serengeti appears to be bounded by rainfall below a threshold of c. 675 mm $year^{-1}$.

In the Serengeti, drier portions of the precipitation gradient show large, homogenous patches of grassland that seem to exclude woody vegetation except in low-lying areas where topographic circumstances, such as river drainages, alter local water availability. Sankaran et al. (2005) purposely excluded low-lying areas and topographic depressions, whereas topography is handled explicitly in the SEM presented above. In the Serengeti the transition from grassland to woodland occurs in many areas between 600 and 700 mm $year^{-1}$ annual rainfall. There is an especially strong transition along the
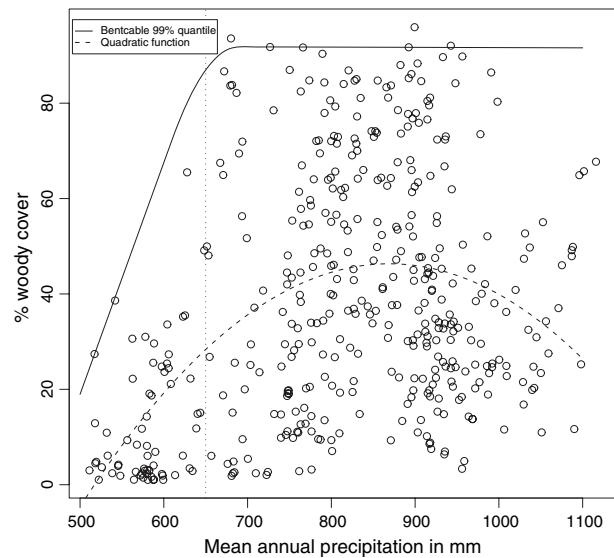

Figure 5 Bivariate plot of the percentage of woody land cover in each Gerresheim subregion (using data from Fig. 2b) as a function of average mean annual rainfall in the same Gerresheim subregion (using data from Fig. 2a). The dashed curve is a quadratic regression of vegetation class values on precipitation. The regression formula is: percentage woody cover = 0.022 × rainfall − 0.000013 × rainfall^2 − 1.83, adjusted R^2 = 0.14. All parameters are significant at the 0.01 level. The solid line is a bent-cable 99% quantile regression. The dashed vertical line marks 650 mm $year^{-1}$ average rainfall, the point below which Sankaran et al. (2005) suggested that woody cover in African savannas is bounded by rainfall.

650 mm $year^{-1}$ rainfall isohyet heading south from the local minima up the slopes of the Ngorongoro highlands, and this appears as a surge of woodland values in the Gerresheim subregions with rainfall values near 650 mm $year^{-1}$ (Fig. 5). It should be noted that the current study uses data that overlap with the Sankaran et al. study, so the two are not entirely independent.

The SEM analysis shows a peak of vegetation heterogeneity at c. 900 mm $year^{-1}$ rainfall, and this coincides with the area between Seronera and Lobo. At the northernmost and westernmost areas of the park, where annual rainfall is highest, there are large patches of tall grassland, although these patches are not as extensive as on the southern plains. The northern and western areas with the largest grassland patches are at the edge of the western corridor and in the Narok Region in Kenya, north and east of the Maasai Mara National Reserve. In both areas, fire may play a critical role (Dempewolf et al., 2007), as may human impact (Serneels & Lambin, 2001).

Increasing human population densities around the national park and protected areas have marked effects on the vegetation. The dense montane forests on the shoulders of the Ngorongoro caldera are clearly visible in the source imagery and the resulting vegetation map, as are the sharp boundaries at the southern edge of NCA, where the forests have been cleared for cultivation. The map also shows areas on the

eastern slopes of NCA where patches of woodland have been cleared for small-scale cultivation within the protected areas. Similar human-induced vegetation is seen along the boundaries of the Maswa Game Control Area in the south-western portion of the ecosystem. The documentation of these patterns in this map will enable monitoring of future human impacts on vegetation in the Serengeti ecosystem.

CONCLUSIONS

The goal of this study was to describe accurately the distribution of major physiognomic vegetation types in the Serengeti ecosystem and to test in what way the spatial configuration of vegetation patches is related to two important ecological factors, namely rainfall and topographic variation. The vegetation map and analysis thereof resulted in three significant causal explanatory models that demonstrate that both rainfall and topography are important contributors to the distribution of woodlands and grassland in the Serengeti, but that these factors alone explain only a fraction of the total variance in patch characteristics and patch diversity throughout the ecosystem. These findings further indicate that patch characteristics do not vary monotonically along the ecological gradient, but rather that the diversity of vegetation types has a complex interaction with rainfall and topography. Furthermore, these findings support the idea that rainfall acts to cap woody tree growth in arid portions of the rainfall gradient (below *c.* 650 mm); above this threshold, other factors such as fire return rates and herbivory become more influential (Sankaran *et al.*, 2005).

Producing and distributing these data in digital format will assist related studies that require detailed information on vegetation structure and distribution in the Serengeti ecosystem. We intend that this product be used to facilitate the conservation and management of the Serengeti ecosystem (e.g. Sinclair *et al.*, 2008) and, more generally, to enhance basic understanding of savanna vegetation in one of the most important and cherished conservation areas on Earth.

ACKNOWLEDGEMENTS

REFERENCES

Anderson, T.M., McNaughton, S.J. & Ritchie, M. (2004) Scale-dependent relationships between the spatial distribution of a limited resource and plant species diversity in an African grassland. *Oecologia*, **139**, 227–287.

Anderson, T.M., Metzger, K.L. & McNaughton, S.J. (2007) Multi-scale analysis of plant species richness in Serengeti grasslands. *Journal of Biogeography*, **34**, 313–323.

Anderson, T.M., Dempewolf, J., Metzger, K.L., Reed, D.N. & Serneels, S. (2008) Generation and maintenance of heterogeneity in the Serengeti ecosystem. *Serengeti III: human impacts on ecosystem dynamics* (ed. by A.R.E. Sinclair, C. Packer, S.A.R. Mduma and J.M. Fryxell), pp. 135–182. University of Chicago Press, Chicago.

Arbuckle, J.L. (2007) *Amos 16.0 users guide*. Amos Development Corporation, Spring House, PA.

Bamford, M.K., Stanistreet, I.G., Stollhofen, C.H. & Albert, R.M. (2008) Late Pliocene grasslands from Olduvai Gorge. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **257**, 280–293.

Belsky, A.J. (1990) Tree/grass ratios in East African savannas: a comparison of existing models. *Journal of Biogeography*, **17**, 483–489.

Belsky, A.J. (1995) Spatial and temporal landscape patterns in arid and semi-arid African savannas. *Mosaic landscapes and ecological processes* (ed. by L. Hansson, L. Fahrig and G. Merriam), pp. 31–56. Chapman & Hall, New York.

Beven, K. & Kirkby, M. (1979) A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, **24**, 43–69.

Burnett, M., August, P.V., Brown, J.H. & Killingbeck, K.T. (1998) The influence of geomorphological heterogeneity on biodiversity I. A patch-scale perspective. *Conservation Biology*, **12**, 363–370.

Chiu, G., Lockhart, R. & Routledge, R. (2006) Bent-cable regression theory and applications. *Journal of the American Statistical Association*, **101**, 542–553.

Cole, M. (1986) *The savannas, biogeography and geobotany*. Academic Press, London.

Coughenour, M.B. (2006) *PPTMAP software*. Fort Collins, CO.

Coughenour, M.B. & Ellis, J.E. (1993) Landscape and climatic control of woody vegetation in a dry tropical ecosystem: Turkana District, Kenya. *Journal of Biogeography*, **20**, 383–398.

Dawson, J.B. (1964) Carbonatitic volcanic ashes in northern Tanganyika. *Bulletin of Volcanology*, **27**, 81–91.

Dempewolf, J., Trigg, S., DeFries, R.S. & Eby, S. (2007) Burned-area mapping of the Serengeti–Mara region using MODIS reflectance data. *IEEE Geoscience and Remote Sensing Letters*, **4**, 312–316.

Field, R., O'Brien, E.M. & Whittaker, R.J. (2005) Global models for predicting woody plant richness from climate: development and evaluation. *Ecology*, **86**, 2263–2277.

Fryxell, J.M., Wilmshurst, J.F., Sinclair, A.R.E., Haydon, D.T., Holt, R.D. & Abrams, P.A. (2005) Landscape scale, heterogeneity, and the viability of Serengeti grazers. *Ecology Letters*, **8**, 328–335.

Gereta, E., Mwangomo, E. & Wolanski, E. (2004) The influence of wetlands in regulating water quality in the Seronera River, Serengeti National Park, Tanzania. *Wetland Ecology and Management*, **12**, 301–307.

Gerresheim, K. (1971) *Serengeti ecosystem landscape classification units*. Tanzania Litho Ltd, Arusha.

Grace, J.B. (2006) *Structural equation modeling and natural systems*. Cambridge University Press, Cambridge.

Grace, J.B. & Bollen, K.A. (2008) Representing general theoretical concepts in structural equation models: the role of composite variables. *Environmental and Ecological Statistics*, **15**, 191–213.

Grunblatt, J., Ottichilo, W.K. & Sinange, R.K. (1989) A hierarchical approach to vegetation classification in Kenya. *African Journal of Ecology*, **27**, 45–51.

Hay, R.L. (1976) *The geology of the Olduvai Gorge*. University of California Press, Berkeley, CA.

Herlocker, D. (1974) *Map of the woody vegetation of the Serengeti National Park*. Texas A&M University, College Station, TX.

Herlocker, D.J., Dirschl, H.J. & Frame, G. (1993) Grasslands of East Africa. *Natural grasslands – eastern hemisphere and résumé* (ed. by R.T. Coupland), pp. 221–264. Elsevier, New York.

Hopcraft, J.G.C., Sinclair, A.R.E. & Packer, C. (2005) Planning for success: Serengeti lions seek prey accessibility rather than abundance. *Journal of Animal Ecology*, **74**, 559–566.

House, J.I., Archer, S., Breshears, D.D. & Scholes, R.J. (2003) Conundrums in mixed woody-herbaceous plant systems. *Journal of Biogeography*, **30**, 1763–1777.

Irish, R.R. (2001) *Landsat 7 Science Data User's Handbook, Report 430-15-01-003-0*. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD, http://landsathandbook.gsfc.nasa.gov/handbook.handbook_toc.html.

Jager, T.J. (1982) *Soils of the Serengeti woodlands, Tanzania*. Agricultural University, Wageningen.

Jensen, J.R. (1996) *Introductory digital image processing: a remote sensing perspective*. Prentice Hall, Upper Saddle River, NJ.

Koenker, R. (2008) Quantreg: quantile regression. *R package version 4.17*.

Kolasa, J. & Waltho, N. (1998) A hierarchical view of habitat and its relationship to species abundance. *Ecological scale*

(ed. by D.L. Peterson and V.T. Parker), pp. 55–76. Columbia University Press, New York.

Lamprey, R.H. & Reid, R.S. (2004) Expansion of human settlement in Kenya's Masai Mara: what future for pastoralism and wildlife? *Journal of Biogeography*, **31**, 997–1032.

Lehner, B., Verdin, K. & Jarvis, A. (2006) *Hydrological data and maps based on Shuttle elevation derivatives at multiple scales (HydroSHEDS) - Technical Documentation*. World Wildlife Fund US, Washington, DC.

Leica (2006) *ERDAS field guide*. Leica Geosystems, Atlanta, GA.

Lillesand, T.M., Kiefer, R.W. & Chipman, J.W. (2008) *Remote sensing and image interpretation*. John Wiley & Sons, New York.

Mackey, B.G. (1994) Predicting the spatial distribution of rainforest structural characteristics. *Journal of Vegetation Science*, **5**, 43–54.

McGarigal, K.S., Cushman, S.A., Neel, M.C. & Ene, E. (2002) *FRAGSTATS: spatial pattern analysis program for categorical maps*. University of Massachusetts, Amherst, MA.

McNaughton, S.J., Banyikwa, F.F. & McNaughton, M.M. (1997) Promotion of the cycling of diet-enhancing nutrients by African grazers. *Science*, **278**, 1798–1800.

Nichols, W.F., Killingbeck, K.T. & August, P.V. (1998) The influence of geomorphological heterogeneity on biodiversity II. A landscape perspective. *Conservation Biology*, **12**, 371–379.

Norton-Griffiths, M. (1975) The patterns of rainfall in the Serengeti ecosystem, Tanzania. *East African Wildlife Journal*, **13**, 347–374.

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

O'Neill, R.V., Milne, B.T., Turner, M.G. & Gardner, R.H. (1988) Resource utilisation scales and landscape pattern. *Landscape Ecology*, **2**, 63–69.

Packer, C., Ikanda, D., Kissui, B. & Kushnir, H. (2005) Conservation biology: Lion attacks on humans in Tanzania. *Nature*, **436**, 927–928.

Reed, D.R. (2003) *Micromammal paleoecology: past and present relationships between African small mammals and their habitats*. Interdepartmental Program in Anthropological Sciences, Stony Brook University, Stony Brook, NY.

Ritchie, M. (1998) Scale-dependent foraging and patch choice in fractal environments. *Evolutionary Ecology*, **12**, 309–330.

Ritchie, M. (2002) Competition and coexistence of mobile animals. *Competition and coexistence* (ed. by U. Sommer and B. Worm), pp. 109–131. Springer-Verlag, Berlin.

Sankaran, M., Ratnam, J. & Hanan, N.P. (2004) Tree–grass coexistence in savannas revisited – insights from an examination of assumptions and mechanisms invoked in existing models. *Ecology Letters*, **7**, 480–490.

Sankaran, M., Hanan, N., Scholes, R., Ratnam, J., Augustine, D., Cade, B., Gignoux, J., Higgins, S., Le Roux, X., Ludwig, F., Ardo, J., Banyikwa, F., Bronn, A., Bucini, G., Caylor, K., Coughenour, M., Diouf, A., Ekaya, W., Feral, C., February,

E., Frost, P., Hiernaux, P., Hrabar, H., Metzger, K., Prins, H., Ringrose, S., Sea, W., Tews, J., Worden, J. & Zambatis, N. (2005) Determinants of woody cover in African savannas. *Nature*, **438**, 846–869.

Scholes, R.D. & Archer, S.R. (1997) Tree–grass interactions in savannas. *Annual Review of Ecology and Systematics*, **28**, 517–544.

Serneels, S. & Lambin, E.F. (2001) Impact of land-use changes on the wildebeest migration in the northern part of the Serengeti–Mara ecosystem. *Journal of Biogeography*, **28**, 391–407.

Sinclair, A.R.E. (1995) Serengeti past and present. *Serengeti II: dynamics, management and conservation of an ecosystem* (ed. by A.R.E. Sinclair and P. Arcese), pp. 3–30. University of Chicago Press, Chicago.

Sinclair, A.R.E., Hopcraft, J.G.C., Olff, H., Mduma, S.A.R., Galvin, K.A. & Sharam, G.J. (2008) Historical and future changes to the Serengeti ecosystem. *Serengeti III: human impacts on ecosystem dynamics* (ed. by A.R.E. Sinclair, C. Packer, S.A.R. Mduma and J.M. Fryxell). Chicago University Press, Chicago.

Solbrig, O. (1996) The diversity of the savanna ecosystem. *Ecological Studies*, **121**, 1–30.

Stohlgren, T.J., Falkner, M.B. & Schell, L.D. (1995) A modified-Whittaker nested vegetation sampling method. *Vegetatio*, **117**, 113–121.

Talbot, L.M. & Kesel, R.H. (1975) The tropical savanna ecosystem. *Geoscience and Man*, **10**, 15–26.

Thirgood, S., Mosser, A., Tham, S., Hopcraft, G., Mwangomo, E., Mlengeya, T., Kilewo, M., Fryxell, J., Sinclair, A.R.E. & Borner, M. (2004) Can parks protect migratory ungulates? The case of the Serengeti wildebeest. *Animal Conservation*, **7**, 113–120.

Urban, D. & Keitt, T. (2001) Landscape connectivity: a graph theory perspective. *Ecology*, **82**, 1205–1218.

Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453–470.

Wilson, J.P. & Gallant, J.C., eds (2000) Digital terrain analysis. *Terrain analysis: principles and applications*. John Wiley & Sons, New York.

deWit, H.A. (1978) *Soils and grassland types of the Serengeti Plain (Tanzania): their distribution and interrelations*. Agricultural University, Wageningen, The Netherlands.

deWit, H.A. & Jeronimus, O.D. (1977) *Soil map of the Serengeti Plains*. Agricultural University, Wageningen, The Netherlands.

Witkowski, E.T. & O'Connor, T.G. (1996) Topo-edaphic, floristic and physiognomic gradients of woody plants in a semi-arid African savanna woodland. *Vegetatio*, **124**, 9–23.

Wolanski, E. & Gereta, E. (2001) Water quantity and quality as the factors driving the Serengeti ecosystem, Tanzania. *Hydrobiologia*, **458**, 169–180.

BIOSKETCH

The Serengeti Mara Working Group comprises ecologists and palaeoecologists researching interrelated topics in the Serengeti–Maasai Mara ecosystem. The workgroup met during a workshop sponsored by the Evolution of Terrestrial Ecosystems Program at the Smithsonian Institution (http://www.nmnh.si.edu/ete/). The workshop was conceived and organized by D.R., a palaeoecologist studying links between savanna ecology and human evolution. He contributed to the remote sensing analysis, vegetation mapping and spatial analysis. T.M.A., a plant ecologist, contributed to the spatial analysis and conducted the SEM analysis. J.D. uses remote sensing to study African fire ecology. He contributed to the remote sensing analysis and vegetation mapping, along with S.S., who studies human impacts on African ecosystems. K.M. is a spatial ecologist. She contributed to the remote sensing analysis and generated the precipitation map. All authors made contributions to the data collection and text of the manuscript.

Editor: David Bowman